

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 21:34:44 ; Search time 36 Seconds

(without alignments)
151.758 Million cell updates/sec

Title: US-09-828-456-2

Perfect score: 207

Sequence: 1 MNRLIELTGMIVLVSVILL.....VASHDNYQPEQASVQHK 41

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing First 45 summaries

Database :

A_Geneseq.101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 207 | 100.0 | 41 | 23 | Escherichia coli b |
| 2 | 60.5 | 29.2 | 231 | 23 | Staphylococcus epi |
| 3 | 58.5 | 28.3 | 203 | 22 | S. epidermidis ope |
| 4 | 58.5 | 28.3 | 203 | 23 | Staphylococcus epi |
| 5 | 58 | 28.0 | 291 | 21 | Staphylococcus epi |
| 6 | 58 | 28.0 | 344 | 21 | Arabidopsis thalia |
| 7 | 58 | 28.0 | 361 | 21 | Arabidopsis thalia |
| 8 | 57 | 27.5 | 327 | 22 | Propionibacterium |
| 9 | 55 | 26.6 | 320 | 23 | Lactococcus lactis |
| 10 | 54.5 | 26.3 | 900 | 22 | C glutamicum prote |

| | | | | | |
|----|------|------|------|----|-----------|
| 11 | 54 | 26.1 | 1048 | 22 | ABG20939 |
| 12 | 53.5 | 25.8 | 206 | 21 | AA11050 |
| 13 | 53.5 | 25.8 | 206 | 22 | AAU34255 |
| 14 | 53.5 | 25.8 | 206 | 22 | AAU37083 |
| 15 | 52 | 25.1 | 455 | 22 | AAU64449 |
| 16 | 52 | 25.1 | 459 | 22 | ABG67668 |
| 17 | 51.5 | 24.9 | 294 | 22 | ABG67023 |
| 18 | 51 | 24.6 | 86 | 20 | AA134616 |
| 19 | 51 | 24.6 | 404 | 23 | ABG95184 |
| 20 | 51 | 24.6 | 541 | 23 | ABG49247 |
| 21 | 51 | 24.6 | 567 | 21 | AA124967 |
| 22 | 51 | 24.6 | 799 | 21 | AA124966 |
| 23 | 51 | 24.6 | 817 | 21 | AA124965 |
| 24 | 51 | 24.6 | 1360 | 22 | ABG61753 |
| 25 | 50.5 | 24.4 | 355 | 22 | ABG57739 |
| 26 | 50.5 | 24.4 | 491 | 23 | ABG38665 |
| 27 | 50 | 24.2 | 232 | 21 | AA138330 |
| 28 | 50 | 24.2 | 248 | 22 | AAU30985 |
| 29 | 50 | 24.2 | 265 | 11 | AA106909 |
| 30 | 50 | 24.2 | 265 | 18 | AA114464 |
| 31 | 50 | 24.2 | 578 | 22 | ABG57874 |
| 32 | 49.5 | 23.9 | 97 | 22 | AAU49588 |
| 33 | 49.5 | 23.9 | 240 | 23 | ABG92584 |
| 34 | 49.5 | 23.9 | 551 | 23 | ABG29598 |
| 35 | 49.5 | 23.9 | 909 | 22 | AAU03554 |
| 36 | 49.5 | 23.9 | 1024 | 23 | ABG07504 |
| 37 | 49 | 23.7 | 103 | 21 | AA105689 |
| 38 | 49 | 23.7 | 137 | 21 | AA105688 |
| 39 | 49 | 23.7 | 146 | 21 | AA105687 |
| 40 | 49 | 23.7 | 146 | 23 | ABG91722 |
| 41 | 49 | 23.7 | 154 | 22 | ABG62506 |
| 42 | 49 | 23.7 | 162 | 21 | AA1033210 |
| 43 | 49 | 23.7 | 176 | 21 | AA1033210 |
| 44 | 49 | 23.7 | 258 | 21 | AA1033209 |
| 45 | 49 | 23.7 | 390 | 11 | AA105475 |

ALIGNMENTS

| | | |
|----------|--|------------------------------|
| RESULT 1 | AA105475 | standard; protein; 41 AA. |
| ID | AA105475 | |
| AC | AA105475 | |
| DT | 08-FEB-2002 (first entry) | |
| DE | Escherichia coli beta-lactam resistance (BLR) protein. | |
| XX | Beta-lactam resistance; blr gene; Beta Lactam-358; BLR; transmembrane; | |
| KW | efflux pump; periplasm; peptidoglycan synthesis inhibition; | |
| KW | drug discovery; screening; antibiotic; antibacterial; vaccine; | |
| KW | gene therapy. | |
| OS | Escherichia coli. | |
| XX | | |
| XX | | |
| FT | Key | Location/Qualifiers |
| FT | Domain | 6..26 |
| XX | | /label= Transmembrane_domain |
| XX | | |
| XX | MO200177176-A2. | |
| XX | | |
| XX | 18-OCT-2001. | |
| XX | | |
| XX | 06-APR-2001; 2001MO-US11363. | |
| XX | | |
| XX | 06-APR-2000; 2000US-195505P. | |
| XX | | |
| XX | 14-JUL-2000; 2000US-218380P. | |
| XX | | |
| XX | (TUFT) TUFTS COLLEGE. | |
| XX | | |
| XX | Levy SB, McMurtry LM; | |
| XX | | |

Novel human diago
Staphylococcus aur
Staphylococcus aur
Staphylococcus aur
Propionibacterium
Drosophila melanog
Drosophila melanog
Chlamydia pneumonia
Human 158P14 sp1
Listeria monocytog
Plant SDF encoded
Plant SDF encoded
Plant SDF encoded
Drosophila melanog
Drosophila melanog
Staphylococcus epi
Human secreted pro
Novel human secret
Intra-acrosomal sp
Acrosomal sperm an
Drosophila melanog
Propionibacterium
Herbicidally activ
Streptococcus poly
Human protein kin
Human GTP-binding
Arabidopsis thalia
Zea mays protein f
Arabidopsis thalia
Herbicidally activ
Drosophila melanog
Zea mays protein f
Arabidopsis thalia
Zea mays protein f
Chloramphenicol re

XX WPI: 2002-010898/01.
DR N-PSDB; ABA02070.
XX
PT New membrane protein, designated Beta Lactam-358 polypeptides, that
PT affect susceptibility to antibiotics which affect peptidoglycan
PT synthesis in microbes, useful for identifying modulators for treating
PT infections
XX
PS Claim 11; Fig 1A; 104pp; English.

XX This sequence represents a beta-lactam resistance (BLR) protein from
CC Escherichia coli. The BLR protein is a 41 amino acid transmembrane
CC protein (also designated Beta Lactam-358 in the specification) with a
CC molecular weight of 4556 daltons which confers resistance to beta-lactam
CC antibiotics. The blr gene which encodes it was initially identified in a
CC 602 bp hypothetical intergenic region of the genome of an antibiotic-
CC susceptible strain of E. coli, RM583. In the RM583 strain, the blr gene
CC is disrupted by insertion of the transposon mlt-Tnp^{hoA}. A clone bearing
CC an ORF (open reading frame)-containing 358 bp segment of the "intergenic"
CC region was found to restore beta-lactam resistance in RM358. Certain
CC amber mutations introduced into the 358 bp clone were found to abolish
CC beta-lactam resistance and resistance could be restored using an
CC amber suppressor, indicating that a protein encoded by the clone was
CC responsible for mediating beta-lactam resistance. It is suggested that
CC the BLR protein may act by increasing a beta-lactamase activity in a
CC cell, possibly being a component of an uncharacterised membrane-bound
CC efflux pump capable of expelling beta-lactams from the periplasm. The
CC BLR protein may be used in methods to identify compounds that modulate
CC antibiotic resistance in a bacterium. The BLR protein, gene, BLR protein
CC homologues, BLR agonists or antagonists, and anti-BLR antibodies may be
CC used to treat bacterial infection, particularly by an organism resistant
CC to antibiotics that affect peptidoglycan synthesis. They may also be
CC used in screening assays, in vaccines, and in diagnostic assays. BLR
CC nucleotides are useful in gene therapy applications to modulate BLR
CC expression, and to detect BLR expression in a biological sample.

SQ Sequence 41 AA:

Query Match 100.0%; Score 207; DB 23; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNRIELTGWIVLVSVILGVASHIDNYQPEQSASVQHK 41
DB 1 MNRIELTGWIVLVSVILGVASHIDNYQPEQSASVQHK 41

RESULT 2
ABP38752
ID ABP38752 standard; Protein; 231 AA.

AC ABP38752;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3597.

KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KM antibacterial; gene therapy.

OS Staphylococcus epidermidis.

PN US680370-B1.

PD 30-APR-2002.

PF 13-AUG-1998; 98US-0134001.

PR 14-AUG-1997; 97US-055779P.

PA 08-NOV-1997; 97US-064964P.

(GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm 1A, Bush D;
XX
XX WPI: 2002-381255/41.
DR N-PSDB; ABN91297.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 3597; 267pp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.

SQ Sequence 231 AA;

Query Match 29.2%; Score 60.5; DB 23; Length 231;
Best Local Similarity 35.0%; Pred. No. 2.5;
Matches 14; Conservative 12; Mismatches 7; Indels 7; Gaps 2;

OY 9 GVIYLVSVILGVASVILGVASHIDNYQPEQSASVQHK 41
DB 8 GIIIFIISVITITGVSALRDKSHEDROKRPQKTSNEQHQ 47

RESULT 3
AAG82688
ID AAG82688 standard; Protein; 203 AA.

AC AAG82688;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2470.

KM Staphylococcus epidermidis SRI strain; infection; diagnosis;
KM vaccination; endocarditis.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

PA (GLAXO) GLAXO GROUP LTD.

PI kimerly WJ;

DR WPI: 2001-316495/33.

DR N-PSDB; AAH53538.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 18; Page 651; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the

CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH5091 to
CC AAH5098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
CC
XX Sequence 203 AA;
SQ
Query Match 28.3%; Score 58.5; DB 22; Length 203;
Best Local Similarity 48.3%; Pred. No. 4.1;
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;
OY 1 MNRLIELTGMIVLWVSIVLLGVASHIDNY 29
||||| | : : : : | : |||||
DB 5 MNRLITLIG-VLLITLAIYLFSPKPYIDNY 32
RESULT 4
ABP39600 standard; Protein; 203 AA.
XX
AC ABP39600;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4445.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-0134001.
XX
PR 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPT; 2002-381255/41.
DR N-PSDB; ABN92145.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 4445; 267Pp; English.
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
XX
SQ Sequence 203 AA;
SQ
Query Match 28.3%; Score 58.5; DB 23; Length 203;
Best Local Similarity 48.3%; Pred. No. 4.1;
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;
OY 1 MNRLIELTGMIVLWVSIVLLGVASHIDNY 29
||||| | : : : : | : |||||
DB 5 MNRLITLIG-VLLITLAIYLFSPKPYIDNY 32
RESULT 5
AAG09408 standard; Protein; 291 AA.
XX
AC AAG09408;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7330.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145952.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150584.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity

28.0%; Score 58; DB 21; Length 291;
34.5%; Pred. No. 7.3;

Matches 10; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

QY 3 RLIEILGWIVVSVLLGASHIDNYP 31
||:| |:::|:| | : : |
DB 212 RLILLGLMLIMGVLLVTSRLIRHP 240

RESULT 6
AAC09407
ID AAC09407 standard; Protein; 344 AA.
XX
AC AAC09407;
XX
XX 17-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7329.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI03405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132867.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137524.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

```
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157453.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158869.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161892.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match      28.0%; Score 58; DB 21; Length 344;
Best Local Similarity 34.5%; Pred. No. 8, 9;
Matches 10; Conservative 10; Mismatches 9; Indels 0; Gaps 0;
```

```
RESULT 7
AAG09406
ID AAG09406 standard; Protein; 361 AA.
XX
AC AAG09406;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7328.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136592.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
```

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144881.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.0%; Score 58; DB 21; Length 361;
Best Local Similarity 34.5%; Pred. No. 9.4;
Matches 10; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

OY 3 RLIELTGIVLVSVLLGVASHIDNTP 31
DB 282 RLILLGLMLIMGVLLVTSRLIRHP 310

RESULT 8
ID AAU53139 standard; Protein; 327 AA.
XX AAU53139;
AC
XX
DT 27-FEB-2002 (first entry)

```

XX DE Propionibacterium acnes immunogenic protein #14035.
XX KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-208841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR WPI: 2001-616774/71.
XX DR N-PsDB: AAS59558.
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX PS Example 1; SEQ ID No 14334; 1069pp; English.
XX CC Sequences AU039105-AAU6017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA).
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 327 AA:
XX
XX Query Match 27.5%; Score 57; DB 22; Length 327;
XX Best Local Similarity 45.5%; Pred. No. 12;
XX Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 7 LTGMIVLVSVLLGVASHIDN 28
DB 249 LAGMLVIVTGGLEAGVAHAHYNN 270

```

```

XX DE Lactococcus lactis protein yqbf.
XX KM Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX OS Lactococcus lactis IL1403.
XX PN FR2807446-A1.
XX PD 12-OCT-2001.
XX PF 11-APR-2000; 2000FR-0004630.
XX PR 11-APR-2000; 2000FR-0004630.
XX PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX PI Bolotine A, Sorokline A, Renault P, Ehrlich SD;
XX PR WPI: 2002-043418/06.
XX PT New nucleotide sequence useful in the identification or Lactococcus
XX PT lactis and related species -
XX PS Claim 6; SEQ ID No 1622; 2504pp; French.
XX CC The present invention is related to a Lactococcus lactis nucleotide
XX CC sequence (ABN90521) and related proteins (ABN53300-ABN55621). The
XX CC nucleic acid sequence is useful in the detection and/or amplification of
XX CC related species. The proteins of the invention are useful for the
XX CC biosynthesis or biodegradation of a composition of interest. The
XX CC invention helps research in lactic bacteria, particularly useful in the
XX CC production of yogurt and cheese.
XX CC Note: The sequence data for this patent is based on equivalent patent
XX CC WO200177334 (published 18-OCT-2001) which is available in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 320 AA:
XX
XX Query Match 26.6%; Score 55; DB 23; Length 320;
XX Best Local Similarity 37.9%; Pred. No. 21;
XX Matches 11; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
QY 1 MNRLIELTGMIVLVSVLLGVASHIDN 29
DB 1 MKRRKMIAMILVLSVLVGGYFAIGNY 29

```

```

RESULT 9
ABB54920
ID ABB54920 standard; Protein: 320 AA.
XX
XX ABB54920;
AC
XX 16-MAY-2002 (first entry)
DT

```

```

RESULT 10
AAG92354
ID AAG92354 standard; Protein: 900 AA.
XX
XX AAG92354;
AC
XX
XX 26-SEP-2001 (first entry)
DE C glutamicum protein fragment SEQ ID NO: 6108.
XX
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX KW organic acid synthesis.
XX OS Corynebacterium glutamicum.
XX PN EPI108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-0127688.
XX PR 16-DEC-1999; 99JP-0377484.
XX PR 07-APR-2000; 2000JP-0159162.
XX PR 03-AUG-2000; 2000JP-0280988.

```

XX (KYOM) KYOMA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
DR WPI: 2001-376931/40.
DR N-PSDB: AAH67573.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 17: SEQ ID NO: 6108; 246pp + Sequence Listing: English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 900 AA;
XX
Query Match 26.3%; Score 54.5; DB 22; Length 900;
Best Local Similarity 34.2%; Pred. No. 84;
Matches 13; Conservative 9; Mismatches 13; Indels 3; Gaps 1;
QY 6 ELTG---WLVVSVLLGVASHIDNYQPEQASVQH 40
||| |||:::|:::|||||:::|
Db 625 ELTDADRWIVDLLEQVRLDVIDALINDNQFAKANDELXH 662
XX
RESULT 11
ABG20939
ID ABG20939 standard; Protein; 1048 AA.
XX
AC ABG20939;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20930.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS85126.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20: SEQ ID NO 51298; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1048 AA;
XX
Query Match 26.1%; Score 54; DB 22; Length 1048;
Best Local Similarity 38.2%; Pred. No. 1.2e+02;
Matches 13; Conservative 9; Mismatches 8; Indels 4; Gaps 2;
QY 11 IVLVSVVLLG---VASHIDNYQPE-QSASVQH 40
:|:::|:::|||||:::|
Db 39 VVVTISVLLGRSIQASRKEKMQPEKEPEVHLH 72
XX
RESULT 12
AAB11050
ID AAB11050 standard; Protein; 206 AA.
XX
AC AAB11050;
XX
DT 16-FEB-2001 (first entry)
XX
DE Staphylococcus aureus sortase transamidase srta protein.
XX
KW Sortase transamidase; Gram-positive bacterium; covalent cross-link;
KW peptidoglycan; antibacterial; vaccine; treatment; detection; srta;
KW bacterial infection.
XX
OS Staphylococcus aureus.
XX
PN WO200062804-A2.
XX
PD 26-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US10198.
XX
PR 15-APR-1999; 99US-0292437.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Schneewind O, Mazmanian S, Liu G, Ton-That H;
XX
DR WPI: 2000-665197/64.
DR N-PSDB: AAC65625.
XX
PT A substantially purified sortase-transamidase from a Gram-positive
PT bacterium for use in the treatment and detection of Gram-positive
PT bacterial infections -


```

CC      P. acnes is also involved in infections of bone, joints and the central
CC      nervous system, however it is particularly involved in the inflammatory
CC      lesions associated with acne vulgaris. A method for detecting the
CC      presence or absence of P. acnes in a patient comprises contacting a
CC      sample with a binding agent that binds to the proteins of the invention
CC      and determining the amount of bound protein in the sample. The
CC      polypeptides may be used as antigens in the production of antibodies
CC      specific for P. acnes proteins. These antibodies can be used to
CC      downregulate expression and activity of P. acnes polypeptides and
CC      therefore treat P. acnes infections. The antibodies may also be used as
CC      diagnostic agents for determining P. acnes presence, for example, by
CC      enzyme linked immunosorbent assay (ELISA).
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence      455 AA;

Query Match      25.1%; Score 52; DB 22; Length 455;
Best Local Similarity 32.6%; Pred. NO. 85;
Matches 14; Conservative 8; Mismatches 13; Indels 8; Gaps 2.

QY      6 ELTGM-----TVLVSVLLLGVAISHINDKQPEQ---SASVQH 40
      ||:::||||| |::||| |
Db      163 EMETFWAMIKIVTIIALLVGGYIAITFGFPPRGSGVPASFSH 205

Search completed: February 6, 2003, 21:38:14
Job time : 38 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: February 6, 2003, 21:37:40 ; Search time 15 Seconds
(without alignment)

(without alignments)

Title: US-09-828-456-2

Sequence: 1 MNRLIELTGWIVLVSVILL.....VASHIDNYQPEQSASVQHK 41

Scoring table: BLOSUM62

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_AA:*

```

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pcp: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pcp: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pcp: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pcp: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pcp: *
6: /cgn2_6/ptodata/1/iaa/backfilltest.pcp: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | | Score | Query Match | Length | DB | ID | Description |
|------------|------|-------|-------------|--------|---------------------|--------------------|-------------|
| 1 | 60.5 | 29.2 | 231 | 4 | US-09-134-001C-3597 | Sequence 3597, App | |
| 2 | 58.5 | 28.3 | 203 | 4 | US-09-134-001C-4445 | Sequence 4445, App | |
| 3 | 52 | 25.1 | 285 | 1 | US-08-292-045-10 | Sequence 10, App | |
| 4 | 50.5 | 24.4 | 491 | 4 | US-09-134-001C-3510 | Sequence 3510, App | |
| 5 | 50 | 23.4 | 265 | 1 | US-08-292-045-2 | Sequence 2, App1 | |
| 6 | 48.5 | 23.4 | 575 | 3 | US-08-912-865-2 | Sequence 2, App1 | |
| 7 | 48.5 | 23.4 | 575 | 4 | US-09-510-949-2 | Sequence 2, App1 | |
| 8 | 46 | 22.2 | 309 | 4 | US-09-134-001C-5667 | Sequence 5667, App | |
| 9 | 46 | 22.2 | 548 | 4 | US-09-601-091-2 | Sequence 2, App1 | |
| 10 | 46 | 22.2 | 548 | 4 | US-09-601-091-4 | Sequence 4, App1 | |
| 11 | 46 | 22.2 | 548 | 4 | US-09-398-395A-52 | Sequence 52, App | |
| 12 | 45.5 | 22.0 | 249 | 4 | US-09-134-001C-4442 | Sequence 4442, App | |
| 13 | 45.5 | 22.0 | 362 | 3 | US-09-120-365-74 | Sequence 74, App | |
| 14 | 45.5 | 22.0 | 362 | 4 | US-09-515-039-74 | Sequence 74, App | |
| 15 | 45.5 | 22.0 | 456 | 4 | US-09-134-001C-3771 | Sequence 3771, App | |
| 16 | 45.5 | 22.0 | 642 | 1 | US-08-706-836-2 | Sequence 2, App1 | |
| 17 | 45.5 | 22.0 | 643 | 2 | US-08-616-844-39 | Sequence 39, App | |
| 18 | 45.5 | 22.0 | 643 | 2 | US-08-599-654-39 | Sequence 39, App | |
| 19 | 45.5 | 22.0 | 643 | 3 | US-08-944-868A-39 | Sequence 39, App | |
| 20 | 45.5 | 22.0 | 643 | 3 | US-08-944-423A-39 | Sequence 39, App | |
| 21 | 45.5 | 22.0 | 643 | 3 | US-08-944-196-39 | Sequence 39, App | |
| 22 | 45 | 21.7 | 80 | 3 | US-09-081-320-31 | Sequence 31, App | |
| 23 | 45 | 21.7 | 80 | 4 | US-09-574-141A-31 | Sequence 31, App | |
| 24 | 45 | 21.7 | 80 | 4 | US-09-707-780-31 | Sequence 31, App | |
| 25 | 45 | 21.7 | 236 | 4 | US-09-134-001C-5548 | Sequence 5548, App | |
| 26 | 45 | 21.7 | 691 | 1 | US-08-178-477B-2 | Sequence 2, App1 | |
| 27 | 44.5 | 21.5 | 142 | 3 | US-08-908-443C-53 | Sequence 53, App | |

| | | | | | | |
|----|------|------|-----|---|---------------------|---------------------|
| 45 | 44 | 21.5 | 166 | 3 | US-08-908-643C-51 | Sequence 51, Appl |
| 29 | 44.5 | 21.5 | 287 | 3 | US-08-549-515-10 | Sequence 10, Appl |
| 28 | 44.5 | 21.5 | 322 | 4 | US-09-457-046B-14 | Sequence 14, Appl |
| 30 | 44.5 | 21.5 | 445 | 4 | US-09-457-046B-52 | Sequence 52, Appl |
| 31 | 44.5 | 21.5 | 572 | 2 | US-08-467-968C-4 | Sequence 4, Appl |
| 32 | 44.5 | 21.5 | 572 | 2 | US-08-838-183D-4 | Sequence 4, Appl |
| 33 | 44.5 | 21.5 | 572 | 2 | US-08-852-344D-4 | Sequence 4, Appl |
| 34 | 44.5 | 21.5 | 572 | 3 | US-08-344-639E-4 | Sequence 4, Appl |
| 35 | 44.5 | 21.5 | 572 | 3 | US-08-987-433-3 | Sequence 3, Appl |
| 36 | 44.5 | 21.5 | 572 | 4 | US-08-467-968A-4 | Sequence 4, Appl |
| 37 | 44.5 | 21.5 | 572 | 4 | US-08-467-968A-4 | Sequence 4, Appl |
| 38 | 44.5 | 21.5 | 572 | 4 | US-08-001-554A-4 | Sequence 4, Appl |
| 39 | 44.5 | 21.5 | 572 | 4 | US-09-134-001C-1165 | Sequence 4165, Appl |
| 40 | 44 | 21.3 | 279 | 4 | US-09-134-001C-9900 | Sequence 4900, Ap |
| 41 | 44 | 21.3 | 282 | 4 | US-09-134-001C-1109 | Sequence 4109, Ap |
| 42 | 44 | 21.3 | 458 | 4 | US-08-487-556E-16 | Sequence 16, Appl |
| 43 | 44 | 21.3 | 489 | 2 | US-08-663-566A-6 | Sequence 6, Appl |
| 44 | 44 | 21.3 | 489 | 2 | US-08-023-610-6 | Sequence 6, Appl |

ALIGNMENTS

RESULT 1
ITS-08-134-001C-3E07

; Sequence 3597, Application US/09134001C
Patent No. 6200370

! GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS AUREUS

;; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SE
; SEQ ID NO 359
;

```

; LENGTH: 231
; TYPE: PRT

```

ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3597

Query Match 29.28: Score 60.5: DB 4: Length 231:

| | | | |
|--------------------------|----------------|----------------|------------|
| Exact matches | 25.26% | Score 0.013 | Length 221 |
| Best Local Similarity | 35.08% | Pred. No. 0.31 | |
| Matches 14: Conservative | 12: Mismatches | 7: Indels | 7: Gaps |
| | | | 2 |

00 9 GWIVLWVSYLLGVA-----SHID--NYOPPEOSASVOKH 41

Db 8 GIIFITSVITGVSAIRPKSHEDRONORPKTSNEOHQ 47

100

RESULT 2
US-09-134-001C-4445

Sequence 4445, Application US/09134001C
Patent No. 6380370

APPLICANT: Lynn Doucette-Stamm et al

APPLICANT: LYNN DOUGLASS-STAMM ET AL
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

1. TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
2. FILE REFERENCE: GTC-007
3. CURRENT APPLICATION NUMBER: US 200 4134 001C

```

; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 00/054,054

```

;; PRIOR APPLICATION NUMBER: US 60/064,964
;;
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/064,964

;; PRIOR APPLICATION NUMBER: US
;;
; PRIOR FILING DATE: 1997-08-14

```

; NUMBER OF SEQ
; SEQ ID NO 4445

```

```

; LENGTH: 2
; TYPE: PRT

```

ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4445

| | | | | |
|-----------------------|-----------------|-----------------|-----------|-------------|
| Query Match | 28.3% | Score 58.5; | DB 4; | Length 203; |
| Best Local Similarity | 48.3%; | Pred. No. 0.52; | | |
| Matches 14; | Conservative 5; | Mismatches 9; | Indels 1; | Gaps 1; |

```
QY      1 MNRLLIELTGWLIVSVSILLGVASHIDNY 29
          ||||| | : : : | | : ||||
Db      5 MNRLLITLIG-VLLILAIYLFESKPYIDNY 32
```

RESULT 3
US-08-292-045-10

GENERAL INFORMATION:
APPLICANT: HERR, JOHN C.
APPLICANT: WRIGHT, RICHARD M.
TITLE OF INVENTION: PRIVATE INTRA-ACROSOMAL SPERM ANTIGEN
TITLE OF INVENTION: FOR USE IN A CONTRACEPTIVE VACCINE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIYAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.

```

?      COMPUTER READABLE FORM.:
?      MEDIUM TYPE: Floppy disk
?      COMPUTER: IBM PC compatible
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: PatentIn Release #1.0, Version #1.25
?      CURRENT APPLICATION DATA:

```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/318,551
FILING DATE: 03-MAR-1989
APPLICATION NUMBER: US 07/481,491
FILING DATE: 16-FEB-1990
APPLICATION NUMBER: US 07/858,798
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 560205man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 494-109-0 C1H
TELECOMMUNICATION INFORMATION:

| | | | | |
|-----------------------|--------|----------------|-------|----------------|
| Query Match | 25.1%; | Score 52; | DB 1; | Length 285; |
| Best Local Similarity | 36.6%; | Pred. No. 7.1; | | |
| Matches | 15; | Conservative | 7; | Mismatches 11; |
| | | | | Indels |

```

Qy      1 MNRLEITGWIIVLVSVIIIGCVASHIDNQPPEDSASVQHK 41
      ||| : ||| : ||| : ||| : ||| :
Db      1 MNRFL-----LMSLYLLGSARGTSS-QPNESGSDIHQ 37

```

RESULT 4
US-09-134-001C-3510

```

; Sequence 3510, Application US/091340010
; Patent No. 6380370
; GENERAL INFORMATION:

```

APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055,779
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 3510
 SEQ ID NO 401

ORGANISM: *Staphylococcus epidermidis*
US-09-134-001C-3510

| | | | | |
|--------------------------|--------|---------------|------------|-------------|
| Query Match | 24.4% | Score 50.5; | DB 4; | Length 491; |
| Best Local Similarity | 31.7%; | Pred. No. 23; | | |
| Matches 13; Conservative | 9; | Mismatches 8; | Indels 11; | Gaps 2. |

```

OY      12 VLAVSVLLGVAS-----HIDNYPPEQGSAS----YQHK 411
      ::||:::| | | | | | | | | | | | | | | | | | | | | |
Db      29 ILVASMFLIGGSVHAAESNHNLENOQSEVSASHSIEMQHK 699

```

RESULT 5
US-08-292-045-2
; Sequence 2, Application US/08292045
; Patent No. 5602005

GENERAL INFORMATION: 1
APPLICANT: HERR, JOHN C. 2
APPLICANT: WRIGHT, RICHARD M. 3
TITLE OF INVENTION: PRIVATE INTRA-ACROSOMAL SPERM ANTIGEN 4
TITLE OF INVENTION: FOR USE IN A CONTRACEPTIVE VACCINE 5
NUMBER OF SEQUENCES: 29 6
CORRESPONDENCE ADDRESS: 7

STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/318,551
FILING DATE: 03-MAR-1989
APPLICATION NUMBER: US 07/481,491
FILING DATE: 16-FEB-1990
APPLICATION NUMBER: US 07/858,798

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

;      LENGTH: 265 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-292-045-2

```

| | | | | |
|--------------------------|-------|----------------|-----------|-------------|
| Query Match | 24.2% | Score 50: | DB 1; | Length 265; |
| Best Local Similarity | 36.6% | Pred. No. 13; | | |
| Matches 15; Conservative | 7; | Mismatches 11; | Indels 8; | Gaps 2 |

```

Qy      1 MNRFLTGWIVLVSVLLGVASHIDNYQPPQASAYQHK 41
      |||: -::|||: |||: |||: |||:
Db      1 MNRFL-----LMSLYLLGSGRGTSS-QPNELSGSIDHQ 33

```

RESULT 6
US-08-922-865-2

```

: GENERAL INFORMATION:
:
: APPLICANT:
:
: TITLE OF INVENTION: MICROORGANISM, LACTAMASE ENZYME OBTAINED
:
: TITLE OF INVENTION: THEREFROM, AND THEIR USE
:
: NUMBER OF SEQUENCES: 2
:
: COMPUTER READABLE FORM:
:

```

```

: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/922,865
: FILING DATE:

```

```

; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-922-865-2

```

| | Query Match | 23.4% | Score 48.5 | DB 3 | length 575 |
|----|---|-----------------|---------------|----------|------------|
| | Best Local Similarity | 34.2% | Pred. No. 55 | | |
| | Matches 13 | Conservative 10 | Mismatches 12 | Indels 3 | Gaps 2 |
| Qy | 5 IELTGIVLVSVILGVASH-IDN-XOPPCDASVQ | 39 | | | |
| | : : : : | : : | | | |
| Db | 236 IEMPGVHMVKVSLINGGMAKYGIKNDIFESPMTENYQ | 333 | | | |

RESULT 7

US-09-510-949-2
; Sequence 2, Application US/09510949
; Patent No. 6423522

APPLICANT: MICROORGANISM, LACTAMASE ENZYME OBTAINED
TITLE OF INVENTION: THEREFROM, AND THEIR USE
TITLE OF INVENTION: THEREFROM, AND THEIR USE
NUMBER OF SEQUENCES: 2
CONTINUED PASTED IN:

```

:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/09/510,949
:
:

```

```

;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-09-510-949-2

```

| | | | | |
|--------------------------|-------|---------------|----------|------------|
| Query Match | 23.4% | Score 48.5 | DB 4 | Length 575 |
| Best Local Similarity | 34.2% | Pred. No. 55 | | |
| Matches 13; Conservative | 10 | Mismatches 12 | Indels 3 | Gaps 2 |

QY 5 IELTGKVLVSVILLGVASH-IDN--YPPESQSASVQ 39
||: ||: ||: | :| :| :| :| :| :| :|
Db 296 IEMPGWTHMKVSILIKGMAKYGILNPILFSPMPMTNYQ 333

RESULT 8
US-09-134-001C-5667
; Sequence 5667, Application US/09134001C

APPLICANT: Lyna Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
FIR. REFERENCE: CMC-007

PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,777
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5667
LENGTH: 309

ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5667

| | | | | |
|--------------------------|--------|----------------|-----------|-------------|
| Query Match | 22.2% | Score 46; | DB 4; | Length 309; |
| Best Local Similarity | 27.8%; | Pred. No. 60; | | |
| Matches 10; Conservative | 8; | Mismatches 14; | Indels 4; | Gaps 1; |

```

QY 10 WIVLWVSVLLGV---ASHIDNYQPPEQSASVQHK 41
    ||| ::| | | | | |
DB 20 WIVICALIVLACGQDANHSSNNKDTESDKKYHR 55

```

```

US-09-601-091-2
; Sequence 2, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: COLBY, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum

```

1 CURRENT FILING DATE: 2000-09-19
2 PRIOR APPLICATION NUMBER: PCT/US99/02133
3 PRIOR FILING DATE: 1999-02-02
4 PRIOR APPLICATION NUMBER: US 60/073,579
5 PRIOR FILING DATE: 1998-02-02

```

1: SEQ ID NO 2
2: LENGTH: 548
3: TYPE: PRF
4: ORGANISM: Lycopersicon esculentum
5: US-09-601-091-2

```

```

Query Match      22.2%   Score 46: DB 4: Length 548;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

QY 11 IVLVSVILLGVASHIDNPQPEQSASVQH 40
   ||| : ||| | | | | | | | | |
Db 61 LVLLIDMQRLLGVAAVHEDN---EIFESIN 86

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
;

```

```

QY      11  IVLVSVILLGVASHIDNYQPPEQASAVQH 40
          ::::  ||||  ||  |  ::::
Db      61  LVLIDAMQRLGVAYHFDN---EIETSIQN 86

```


ORGANISM: Vigna mungo CP
US-09-515-039-74

| | | | | |
|--------------------------|--------|---------------|-----------|-------------|
| Query Match | 22.08; | Score 45.5; | DB 4; | Length 362; |
| Best Local Similarity | 34.38; | Pred. No. 86; | | |
| Matches 12; Conservative | 10; | Mismatches 8; | Indels 5; | Gaps 2; |

```
QY      1  MNRLIELTGMIVLVSVILLGVASHIDNQPPEQS 35
      2  |::|::|::|::|::|::|::|::|::|::|
Db      3  MKKLL---WVVLSTSLV-LGVANSFDFHEKDLES 32
```

```

RESULT 15
US-09-134-001C-3771
: Sequence 3771. Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 3771
: LENGTH: 456
: TYPE: PRT
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3771

```

| | | | | |
|--------------------------|--------|--------------------|-----------|-------------|
| Query Match | 22.0%; | Score 45.5; | DB 4; | Length 456; |
| Best Local Similarity | 27.5%; | Pred. NO. 1.1e+02; | | |
| Matches 11; Conservative | 11; | Mismatches 13; | Indels 5; | Gaps 1; |

```

Qy      6 ELTGW-----IYLVSVAILLGVAASHIDNYQPEQASAYQH 40
          || |      || :::::|| : :: | : |
Db     149 ELEFWPAIKIYVILLVIGVMILFAFKTPRGNTSLTH 188

```

Search completed: February 6, 2003, 21:39:54
Job time : 17 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 21:38:20 : Search time 11 Seconds

(without alignments)
82.631 Million cell updates/sec

Title: US-09-828-456-2

Perfect score: 207

Sequence: 1 MNRLLIETGMIYLVSVILL.....VASHIDNYPQASVQHK 41

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues

Total number of hits satisfying chosen parameters: 129505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications-AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 207 | 100.0 | 41 | 10 | US-09-828-456-2 |
| 2 | 54.5 | 26.3 | 900 | 9 | US-09-738-626-6108 |
| 3 | 53.5 | 25.8 | 206 | 9 | US-09-833-999A-3 |
| 4 | 53.5 | 25.8 | 206 | 10 | US-09-815-242-5751 |
| 5 | 53.5 | 25.8 | 206 | 10 | US-09-815-242-12676 |
| 6 | 49 | 23.7 | 1053 | 10 | US-09-815-242-5136 |
| 7 | 48.5 | 23.4 | 323 | 9 | US-09-978-295A-330 |
| 8 | 48.5 | 23.4 | 323 | 9 | US-09-978-697-330 |
| 9 | 48.5 | 23.4 | 323 | 9 | US-09-978-192A-330 |
| 10 | 48.5 | 23.4 | 323 | 9 | US-09-999-832A-330 |
| 11 | 48.5 | 23.4 | 323 | 9 | US-09-978-189-330 |
| 12 | 48.5 | 23.4 | 323 | 9 | US-10-028-072-272 |
| 13 | 48.5 | 23.4 | 323 | 9 | US-10-121-049-272 |
| 14 | 48.5 | 23.4 | 323 | 9 | US-10-123-904-272 |
| 15 | 48.5 | 23.4 | 323 | 9 | US-10-140-470-272 |
| 16 | 48.5 | 23.4 | 360 | 10 | US-09-815-242-12011 |
| 17 | 48 | 23.2 | 81 | 9 | US-09-738-626-5932 |
| 18 | 48 | 23.2 | 497 | 12 | US-10-033-109-10 |
| 19 | 47 | 22.7 | 178 | 9 | US-09-738-626-5509 |

| | | | | | | |
|----|------|------|-----|----|---------------------|--------------------|
| 20 | 47 | 22.7 | 397 | 10 | US-09-815-242-12699 | Sequence 12699, A |
| 21 | 46.5 | 22.5 | 83 | 10 | US-09-864-761-46530 | Sequence 46530, A |
| 22 | 46.5 | 22.5 | 444 | 10 | US-09-881-752A-212 | Sequence 212, App |
| 23 | 46 | 22.2 | 530 | 9 | US-10-047-542-85 | Sequence 85, App |
| 24 | 46 | 22.2 | 548 | 9 | US-09-738-626-5776 | Sequence 5776, App |
| 25 | 46 | 22.2 | 548 | 10 | US-09-887-586A-52 | Sequence 52, App |
| 26 | 46 | 22.2 | 548 | 10 | US-09-903-012-52 | Sequence 77, App |
| 27 | 45.5 | 22.0 | 44 | 10 | US-09-739-907-77 | Sequence 160, App |
| 28 | 45.5 | 22.0 | 58 | 10 | US-09-739-907-160 | Sequence 639, App |
| 29 | 45.5 | 22.0 | 61 | 10 | US-09-764-847-639 | Sequence 74, App |
| 30 | 45.5 | 22.0 | 537 | 9 | US-10-047-542-74 | Sequence 39, App |
| 31 | 45.5 | 22.0 | 643 | 9 | US-10-060-763-11 | Sequence 56, App |
| 32 | 45.5 | 22.0 | 643 | 10 | US-09-371-900-39 | Sequence 11, App |
| 33 | 45.5 | 22.0 | 643 | 10 | US-09-924-417-56 | Sequence 56, App |
| 34 | 45.5 | 22.0 | 643 | 10 | US-09-924-417-56 | Sequence 55, App |
| 35 | 45 | 21.7 | 94 | 10 | US-09-778-927A-55 | Sequence 1330, App |
| 36 | 45 | 21.7 | 154 | 10 | US-09-867-550-1330 | Sequence 6459, App |
| 37 | 45 | 21.7 | 159 | 9 | US-09-738-626-6459 | Sequence 5178, App |
| 38 | 45 | 21.7 | 162 | 9 | US-09-738-626-5178 | Sequence 43, App |
| 39 | 45 | 21.7 | 299 | 10 | US-09-510-332-9 | Sequence 9, App |
| 40 | 45 | 21.7 | 299 | 10 | US-09-393-634-43 | Sequence 6415, App |
| 41 | 45 | 21.7 | 317 | 9 | US-09-738-626-6415 | Sequence 6555, App |
| 42 | 45 | 21.7 | 449 | 9 | US-09-738-626-6555 | Sequence 28, App |
| 43 | 45 | 21.7 | 464 | 9 | US-10-024-623-28 | Sequence 39, App |
| 44 | 45 | 21.7 | 473 | 10 | US-09-790-264-39 | Sequence 84, App |
| 45 | 45 | 21.7 | 535 | 9 | US-10-047-542-84 | |

ALIGNMENTS

RESULT 1
US-09-828-456-2
Sequence 2, Application US/09828456
Patent No. US20020051982A1
GENERAL INFORMATION:
APPLICANT: Levy, Stewart et al.
TITLE OF INVENTION: NO. US20020051982A1 BLR Molecules Affecting Antibiotic Susce
FILE REFERENCE: PKZ-030
CURRENT APPLICATION NUMBER: US/09/828,456
CURRENT FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/195,505
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/218,380
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 41
TYPE: PRT
ORGANISM: Escherichia coli
US-09-828-456-2

Query Match 100.0%; Score 207; DB 10;
Best Local Similarity 100.0%; Pred No. 1.5e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRLLIETGMIYLVSVILLGVASHIDNYPQASVQHK 41
DB 1 MNRLLIETGMIYLVSVILLGVASHIDNYPQASVQHK 41

RESULT 2
US-09-738-626-6108
Sequence 6108, Application US/09738626
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO

```

; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKARI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 6108
; LENGTH: 900
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-6108

Query Match          26.3%; Score 54.5; DB 9; Length 900;
Best Local Similarity 34.2%; Pred. No. 18;
Matches 13; Conservative 9; Mismatches 13; Indels 3; Gaps 1;

QY 6 ELTG---WYLVVSVILGVASHIDNYQPEQSASVQH 40
      |||  |||  :  |  |  :  |||  :  :  |
DB 625 ELTDADRWIVDLLEQVRLDVADLVNQFAKANEELYH 662

RESULT 3
; US-09-933-999A-3
; Sequence 3, Application US/09933999A
; Publication No. US20030022178A1
; GENERAL INFORMATION:
; APPLICANT: U.C. Regents
; APPLICANT: Schneewind et al.
; TITLE OF INVENTION: Identification of Sortase Gene
; FILE REFERENCE: 510015-249
; CURRENT APPLICATION NUMBER: US/09/933,999A
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 3
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-933-999A-3

Query Match          25.8%; Score 53.5; DB 9; Length 206;
Best Local Similarity 42.9%; Pred. No. 4.4;
Matches 12; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 2 NRLELTGWIWLVSVILGVASHIDNY 29
      |||  :  |  |  :  |  |||  |||
DB 6 NRLEMTAG-VLLIVAVYLFAPKPHIDNY 32

RESULT 4
; US-09-815-242-5751
; Sequence 5751, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
```

```

; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 5751
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-5751

Query Match          25.8%; Score 53.5; DB 10; Length 206;
Best Local Similarity 42.9%; Pred. No. 4.4;
Matches 12; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 2 NRLELTGWIWLVSVILGVASHIDNY 29
      |||  :  |  |  :  |  |||  |||
DB 6 NRLEMTAG-VLLIVAVYLFAPKPHIDNY 32

RESULT 5
; US-09-815-242-12676
; Sequence 12676, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12676
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
```


| | | |
|----|-------------------------------------|--|
| 1 | PRIOR FILING DATE: 1998-03-31 | |
| 2 | PRIOR APPLICATION NUMBER: 60/080107 | |
| 3 | PRIOR FILING DATE: 1998-03-31 | |
| 4 | PRIOR APPLICATION NUMBER: 60/080165 | |
| 5 | PRIOR FILING DATE: 1998-03-31 | |
| 6 | PRIOR APPLICATION NUMBER: 60/080194 | |
| 7 | PRIOR FILING DATE: 1998-03-31 | |
| 8 | PRIOR APPLICATION NUMBER: 60/080327 | |
| 9 | PRIOR FILING DATE: 1998-04-01 | |
| 10 | PRIOR APPLICATION NUMBER: 60/080328 | |
| 11 | PRIOR FILING DATE: 1998-04-01 | |
| 12 | PRIOR APPLICATION NUMBER: 60/080333 | |
| 13 | PRIOR FILING DATE: 1998-04-01 | |
| 14 | PRIOR APPLICATION NUMBER: 60/080334 | |
| 15 | PRIOR FILING DATE: 1998-04-01 | |
| 16 | PRIOR APPLICATION NUMBER: 60/081070 | |
| 17 | PRIOR FILING DATE: 1998-04-08 | |
| 18 | PRIOR APPLICATION NUMBER: 60/081049 | |
| 19 | PRIOR FILING DATE: 1998-04-08 | |
| 20 | PRIOR APPLICATION NUMBER: 60/081071 | |
| 21 | PRIOR FILING DATE: 1998-04-08 | |
| 22 | PRIOR APPLICATION NUMBER: 60/081195 | |
| 23 | PRIOR FILING DATE: 1998-04-08 | |
| 24 | PRIOR APPLICATION NUMBER: 60/081203 | |
| 25 | PRIOR FILING DATE: 1998-04-09 | |
| 26 | PRIOR APPLICATION NUMBER: 60/081229 | |
| 27 | PRIOR FILING DATE: 1998-04-09 | |
| 28 | PRIOR APPLICATION NUMBER: 60/081955 | |
| 29 | PRIOR FILING DATE: 1998-04-15 | |
| 30 | PRIOR APPLICATION NUMBER: 60/081817 | |
| 31 | PRIOR FILING DATE: 1998-04-15 | |
| 32 | PRIOR APPLICATION NUMBER: 60/081819 | |
| 33 | PRIOR FILING DATE: 1998-04-15 | |
| 34 | PRIOR APPLICATION NUMBER: 60/081952 | |
| 35 | PRIOR FILING DATE: 1998-04-15 | |
| 36 | PRIOR APPLICATION NUMBER: 60/081838 | |
| 37 | PRIOR FILING DATE: 1998-04-15 | |
| 38 | PRIOR APPLICATION NUMBER: 60/082568 | |
| 39 | PRIOR FILING DATE: 1998-04-21 | |
| 40 | PRIOR APPLICATION NUMBER: 60/082569 | |
| 41 | PRIOR FILING DATE: 1998-04-21 | |
| 42 | PRIOR APPLICATION NUMBER: 60/082704 | |
| 43 | PRIOR FILING DATE: 1998-04-22 | |
| 44 | PRIOR APPLICATION NUMBER: 60/082804 | |
| 45 | PRIOR FILING DATE: 1998-04-22 | |
| 46 | PRIOR APPLICATION NUMBER: 60/082700 | |
| 47 | PRIOR FILING DATE: 1998-04-22 | |
| 48 | PRIOR APPLICATION NUMBER: 60/082797 | |
| 49 | PRIOR FILING DATE: 1998-04-22 | |
| 50 | PRIOR APPLICATION NUMBER: 60/082796 | |
| 51 | PRIOR FILING DATE: 1998-04-23 | |
| 52 | PRIOR APPLICATION NUMBER: 60/083336 | |
| 53 | PRIOR FILING DATE: 1998-04-27 | |
| 54 | PRIOR APPLICATION NUMBER: 60/083322 | |
| 55 | PRIOR FILING DATE: 1998-04-28 | |
| 56 | PRIOR APPLICATION NUMBER: 60/083392 | |
| 57 | PRIOR FILING DATE: 1998-04-29 | |
| 58 | PRIOR APPLICATION NUMBER: 60/083495 | |
| 59 | PRIOR FILING DATE: 1998-04-29 | |
| 60 | PRIOR APPLICATION NUMBER: 60/083496 | |
| 61 | PRIOR FILING DATE: 1998-04-29 | |
| 62 | PRIOR APPLICATION NUMBER: 60/083499 | |
| 63 | PRIOR FILING DATE: 1998-04-29 | |
| 64 | PRIOR APPLICATION NUMBER: 60/083545 | |
| 65 | PRIOR FILING DATE: 1998-04-29 | |
| 66 | PRIOR APPLICATION NUMBER: 60/083544 | |
| 67 | PRIOR FILING DATE: 1998-04-29 | |
| 68 | PRIOR APPLICATION NUMBER: 60/083558 | |
| 69 | PRIOR FILING DATE: 1998-04-29 | |
| 70 | PRIOR APPLICATION NUMBER: 60/083559 | |
| 71 | PRIOR FILING DATE: 1998-04-29 | |
| 72 | PRIOR APPLICATION NUMBER: 60/083500 | |
| 73 | PRIOR FILING DATE: 1998-04-29 | |

```

PRIORITY APPLICATION NUMBER: 60/083742
PRIORITY FILING DATE: 1998-04-30
PRIORITY APPLICATION NUMBER: 60/084366
PRIORITY FILING DATE: 1998-05-05
PRIORITY APPLICATION NUMBER: 60/084414
PRIORITY FILING DATE: 1998-05-06
PRIORITY APPLICATION NUMBER: 60/084441
PRIORITY FILING DATE: 1998-05-06
PRIORITY APPLICATION NUMBER: 60/084637
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/084639
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/084640
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/084598
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/084600
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/084627
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/084643
PRIORITY FILING DATE: 1998-05-07
PRIORITY APPLICATION NUMBER: 60/085339
PRIORITY FILING DATE: 1998-05-13
PRIORITY APPLICATION NUMBER: 60/085338
PRIORITY FILING DATE: 1998-05-13
PRIORITY APPLICATION NUMBER: 60/085323
PRIORITY FILING DATE: 1998-05-13
PRIORITY APPLICATION NUMBER: 60/085582
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085700
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085689
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085579
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085580
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085573
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085704
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/085697

Query Match 23.4%, Score 48.5, DB 9, Length 323,
Best Local Similarity 36.1%, Pred. No. 35,
Matches 13, Conservative 9, Mismatches 5, Indels 9, Gaps 2

QY 8 TGMIV--LVVSVILL-----GVASHIDNYQPPDQ 34
      |||: |||::| ||::| ||:
Db 238 SGMILLTTLVLSVWLMWICATVATVAYQVPSK 273

RESULT 8
US-09-978-697-330
; Sequence 330, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.

```

APPLICANT: Hillan, Kenneth J
APPLICANT: KJavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/055311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441

PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 23.4% Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1% Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

Qy 8 TGMIV--LVSVYIL-----GVASHIDNYQPPQ 34
Db 238 SCWILTTTLVLSVWLLMICATVATAVEQYVPSK 273

RESULT 9
US-09-978-192A-330
Sequence 330, Application US/09978192A
Patent No. US2002017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann

APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334

```

; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

QY 8 TGWIV---LVSVVLL-----GVASHIDNYQPEQ 34
Db 238 SGWILFTTVLVSVMVLMICCAVFAVAVEQVPSEK 273

RESULT 10
US-09-999-832A-330
; Sequence 330, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashtkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltzen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
```

FILE REFERENCE: P2630PIC63
CURRENT APPLICATION NUMBER: US/09/999,832A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/08336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643

;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

OY 8 TGMIV--LVSVLL-----GVASHIDNYQPEQ 34
DB 238 SGWILTTTLVLSVWVLMICCAVATAVEQYVPSEK 273

RESULT 11

US-09-978-189-330
; Sequence 330, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Peoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC7
; CURRENT APPLICATION NUMBER: US/09/978,189
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

;; PRIOR APPLICATION NUMBER: 60/064249
;; PRIOR FILING DATE: 1997-11-03
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066364
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/077450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077641
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/07791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078004
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078936
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955

PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083455
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13

PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 23.4% Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

OY 8 TGMIV---LVSVYIL-----GVASHIDNYQPPQ 34
DB 238 SGWILTTLVLSYMWLWICATVATVAYQYPSK 273

RESULT 12
US-10-028-072-272
; Sequence 272, Application US/10028072
; Publication No. US2003000431A1

GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang

TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24

| | |
|---------------------------|------------|
| PRIOR APPLICATION NUMBER: | 60/062250 |
| PRIOR FILING DATE: | 1997-10-17 |
| PRIOR APPLICATION NUMBER: | 60/062285 |
| PRIOR FILING DATE: | 1997-10-17 |
| PRIOR APPLICATION NUMBER: | 60/062287 |
| PRIOR FILING DATE: | 1997-10-17 |
| PRIOR APPLICATION NUMBER: | 60/062814 |
| PRIOR FILING DATE: | 1997-10-24 |
| PRIOR APPLICATION NUMBER: | 60/062816 |
| PRIOR FILING DATE: | 1997-10-24 |
| PRIOR APPLICATION NUMBER: | 60/063045 |
| PRIOR FILING DATE: | 1997-10-24 |
| PRIOR APPLICATION NUMBER: | 60/063082 |
| PRIOR FILING DATE: | 1997-10-31 |
| PRIOR APPLICATION NUMBER: | 60/063127 |
| PRIOR FILING DATE: | 1997-10-24 |
| PRIOR APPLICATION NUMBER: | 60/063322 |
| PRIOR FILING DATE: | 1997-10-27 |
| PRIOR APPLICATION NUMBER: | 60/063329 |
| PRIOR FILING DATE: | 1997-10-27 |
| PRIOR APPLICATION NUMBER: | 60/063550 |
| PRIOR FILING DATE: | 1997-10-28 |
| PRIOR APPLICATION NUMBER: | 60/063561 |
| PRIOR FILING DATE: | 1997-10-28 |
| PRIOR APPLICATION NUMBER: | 60/063704 |
| PRIOR FILING DATE: | 1997-10-29 |
| PRIOR APPLICATION NUMBER: | 60/063733 |
| PRIOR FILING DATE: | 1997-10-29 |
| PRIOR APPLICATION NUMBER: | 60/063735 |
| PRIOR FILING DATE: | 1997-10-29 |
| PRIOR APPLICATION NUMBER: | 60/064248 |
| PRIOR FILING DATE: | 1997-11-03 |
| PRIOR APPLICATION NUMBER: | 60/064809 |
| PRIOR FILING DATE: | 1997-11-07 |
| PRIOR APPLICATION NUMBER: | 60/065168 |
| PRIOR FILING DATE: | 1997-11-12 |
| PRIOR APPLICATION NUMBER: | 60/065846 |
| PRIOR FILING DATE: | 1997-11-17 |
| PRIOR APPLICATION NUMBER: | 60/066364 |
| PRIOR FILING DATE: | 1997-11-21 |
| PRIOR APPLICATION NUMBER: | 60/066453 |
| PRIOR FILING DATE: | 1997-11-24 |
| PRIOR APPLICATION NUMBER: | 60/066511 |
| PRIOR FILING DATE: | 1997-11-24 |
| PRIOR APPLICATION NUMBER: | 60/066770 |
| PRIOR FILING DATE: | 1997-11-24 |
| PRIOR APPLICATION NUMBER: | 60/069212 |
| PRIOR FILING DATE: | 1997-12-11 |
| PRIOR APPLICATION NUMBER: | 60/069278 |
| PRIOR FILING DATE: | 1997-12-11 |
| PRIOR APPLICATION NUMBER: | 60/069334 |
| PRIOR FILING DATE: | 1997-12-11 |
| PRIOR APPLICATION NUMBER: | 60/069694 |
| PRIOR FILING DATE: | 1997-12-16 |
| PRIOR APPLICATION NUMBER: | 60/072320 |
| PRIOR FILING DATE: | 1998-01-23 |
| PRIOR APPLICATION NUMBER: | 60/073612 |
| PRIOR FILING DATE: | 1998-07-04 |
| PRIOR APPLICATION NUMBER: | 60/074086 |
| PRIOR FILING DATE: | 1998-07-05 |
| PRIOR APPLICATION NUMBER: | 60/074092 |
| PRIOR FILING DATE: | 1998-07-09 |
| PRIOR APPLICATION NUMBER: | 60/077791 |
| PRIOR FILING DATE: | 1998-03-12 |
| PRIOR APPLICATION NUMBER: | 60/078910 |
| PRIOR FILING DATE: | 1998-03-20 |
| PRIOR APPLICATION NUMBER: | 60/079294 |
| PRIOR FILING DATE: | 1998-03-25 |
| PRIOR APPLICATION NUMBER: | 60/079663 |
| PRIOR FILING DATE: | 1998-03-25 |

| | |
|----|-------------------------------------|
| 1 | PRIOR FILING DATE: 1998-02-27 |
| 2 | PRIOR APPLICATION NUMBER: 60/079728 |
| 3 | PRIOR FILING DATE: 1998-03-27 |
| 4 | PRIOR APPLICATION NUMBER: 60/080165 |
| 5 | PRIOR FILING DATE: 1998-03-31 |
| 6 | PRIOR APPLICATION NUMBER: 60/081203 |
| 7 | PRIOR FILING DATE: 1998-04-09 |
| 8 | PRIOR APPLICATION NUMBER: 60/081229 |
| 9 | PRIOR FILING DATE: 1998-04-09 |
| 10 | PRIOR APPLICATION NUMBER: 60/081655 |
| 11 | PRIOR FILING DATE: 1998-04-14 |
| 12 | PRIOR APPLICATION NUMBER: 60/081817 |
| 13 | PRIOR FILING DATE: 1998-04-15 |
| 14 | PRIOR APPLICATION NUMBER: 60/081818 |
| 15 | PRIOR FILING DATE: 1998-04-15 |
| 16 | PRIOR APPLICATION NUMBER: 60/082999 |
| 17 | PRIOR FILING DATE: 1998-04-22 |
| 18 | PRIOR APPLICATION NUMBER: 60/083545 |
| 19 | PRIOR FILING DATE: 1998-04-24 |
| 20 | PRIOR APPLICATION NUMBER: 60/083322 |
| 21 | PRIOR FILING DATE: 1998-04-28 |
| 22 | PRIOR APPLICATION NUMBER: 60/083545 |
| 23 | PRIOR FILING DATE: 1998-04-29 |
| 24 | PRIOR APPLICATION NUMBER: 60/084600 |
| 25 | PRIOR FILING DATE: 1998-05-07 |
| 26 | PRIOR APPLICATION NUMBER: 60/084627 |
| 27 | PRIOR FILING DATE: 1998-05-07 |
| 28 | PRIOR APPLICATION NUMBER: 60/084637 |
| 29 | PRIOR FILING DATE: 1998-05-07 |
| 30 | PRIOR APPLICATION NUMBER: 60/085338 |
| 31 | PRIOR FILING DATE: 1998-05-12 |
| 32 | PRIOR APPLICATION NUMBER: 60/085332 |
| 33 | PRIOR FILING DATE: 1998-05-13 |
| 34 | PRIOR APPLICATION NUMBER: 60/085338 |
| 35 | PRIOR FILING DATE: 1998-05-13 |
| 36 | PRIOR APPLICATION NUMBER: 60/085339 |
| 37 | PRIOR FILING DATE: 1998-05-13 |
| 38 | PRIOR APPLICATION NUMBER: 60/085579 |
| 39 | PRIOR FILING DATE: 1998-05-15 |
| 40 | PRIOR APPLICATION NUMBER: 60/085637 |
| 41 | PRIOR FILING DATE: 1998-05-15 |
| 42 | PRIOR APPLICATION NUMBER: 60/085704 |
| 43 | PRIOR FILING DATE: 1998-05-15 |
| 44 | PRIOR APPLICATION NUMBER: 60/086414 |
| 45 | PRIOR FILING DATE: 1998-05-22 |
| 46 | PRIOR APPLICATION NUMBER: 60/086430 |
| 47 | PRIOR FILING DATE: 1998-05-22 |
| 48 | PRIOR APPLICATION NUMBER: 60/087106 |
| 49 | PRIOR FILING DATE: 1998-05-28 |
| 50 | PRIOR APPLICATION NUMBER: 60/088026 |
| 51 | PRIOR FILING DATE: 1998-06-04 |
| 52 | PRIOR APPLICATION NUMBER: 60/088730 |
| 53 | PRIOR FILING DATE: 1998-06-10 |
| 54 | PRIOR APPLICATION NUMBER: 60/088741 |
| 55 | PRIOR FILING DATE: 1998-06-10 |
| 56 | PRIOR APPLICATION NUMBER: 60/088810 |
| 57 | PRIOR FILING DATE: 1998-06-17 |
| 58 | PRIOR APPLICATION NUMBER: 60/089599 |
| 59 | PRIOR FILING DATE: 1998-06-17 |
| 60 | PRIOR APPLICATION NUMBER: 60/089907 |
| 61 | PRIOR FILING DATE: 1998-06-18 |
| 62 | PRIOR APPLICATION NUMBER: 60/089974 |
| 63 | PRIOR FILING DATE: 1998-06-19 |
| 64 | PRIOR APPLICATION NUMBER: 60/090349 |
| 65 | PRIOR FILING DATE: 1998-06-23 |
| 66 | PRIOR APPLICATION NUMBER: 60/090429 |
| 67 | PRIOR FILING DATE: 1998-06-24 |
| 68 | PRIOR APPLICATION NUMBER: 60/090445 |
| 69 | PRIOR FILING DATE: 1998-06-24 |
| 70 | PRIOR APPLICATION NUMBER: 60/090538 |
| 71 | PRIOR FILING DATE: 1998-06-24 |

```
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match      23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

OY      8 TGMIV---LVSVYLL-----GVASHIDNYOPPEQ 34
DB      238 SGWLTITTLVLSVWVLMICATVATAVEQYVPSK 273

RESULT 13
US-10-121-049-272
; Sequence 272, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 272
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-121-049-272

Query Match      23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

OY      8 TGMIV---LVSVYLL-----GVASHIDNYOPPEQ 34
DB      238 SGWLTITTLVLSVWVLMICATVATAVEQYVPSK 273

RESULT 14
US-10-123-904-272
; Sequence 272, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
```

```
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 272
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-123-904-272

Query Match      23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

OY      8 TGMIV---LVSVYLL-----GVASHIDNYOPPEQ 34
DB      238 SGWLTITTLVLSVWVLMICATVATAVEQYVPSK 273

RESULT 15
US-10-140-470-272
; Sequence 272, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 272
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo Saplen
US-10-140-470-272

Query Match      23.4%; Score 48.5; DB 9; Length 323;
Best Local Similarity 36.1%; Pred. No. 35;
Matches 13; Conservative 9; Mismatches 5; Indels 9; Gaps 2;

OY      8 TGMIV---LVSVYLL-----GVASHIDNYOPPEQ 34
```

Mon Feb 10 09:26:00 2003

us-09-828-456-2.rapb

Page 13

DB 238 SGWITFTTLYLSVMVLLMICCATVATAVEQYVPSK 273

Search completed: February 6, 2003, 21:40:11
Job time : 12 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 05:05:32 ; Search time 2578 Seconds

(without alignments)
5429.965 Million cell updates/sec

Title: US-09-828-456-1

Sequence: 1 tgcctcctatggaagatgaa.....gtgattgtcaggaagataaa 481

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-----------|---------------------|
| 1 | 481 | 100.0 | 481 | 6 | AX278241 | AX278241 Sequence |
| 2 | 469 | 97.5 | 13172 | 1 | AE000258 | AE000258 Escherich |
| 3 | 469 | 97.5 | 15332 | 1 | D90806 | D90806 E.coli geno |
| 4 | 469 | 97.5 | 15601 | 1 | D90808 | D90808 E.coli geno |
| 5 | 469 | 97.5 | 20231 | 1 | D90807 | D90807 E.coli geno |
| 6 | 452.2 | 94.0 | 13259 | 1 | AE005386 | AE005386 Escherich |
| 7 | 449.8 | 93.5 | 257071 | 1 | AF002558 | AF002558 Escherich |
| 8 | 347 | 72.1 | 358 | 1 | AP219227 | AP219227 Escherich |
| 9 | 203 | 42.2 | 19211 | 1 | D90805 | D90805 E.coli geno |
| 10 | 106.6 | 22.2 | 23009 | 1 | AE008763 | AE008763 Salmonell |
| 11 | 103.4 | 21.5 | 233050 | 1 | AL627271 | AL627271 Salmonell |
| 12 | 42.4 | 8.8 | 134978 | 2 | AC079004 | AC079004 Homo sapi |
| 13 | 42.4 | 8.8 | 164184 | 2 | AC079050 | AC079050 Homo sapi |
| 14 | 42.4 | 8.8 | 167577 | 9 | AC099687 | AC099687 Homo sapi |
| 15 | 42.2 | 8.8 | 190015 | 2 | AL591126 | AL591126 Mus muscu |
| 16 | 41.8 | 8.7 | 67890 | 8 | NCB14H13 | NCB14H13 Neurospor |
| 17 | 41.2 | 8.6 | 175111 | 2 | AC101999 | AC101999 Mus muscu |
| 18 | 41.2 | 8.6 | 204591 | 2 | AC122262 | AC122262 Mus muscu |
| 19 | 41 | 8.5 | 95417 | 9 | AC010876 | AC010876 Homo sapi |
| 20 | 40.6 | 8.4 | 160824 | 2 | AC069118 | AC069118 Homo sapi |
| 21 | 40.6 | 8.4 | 173770 | 2 | AC122716 | AC122716 Homo sapi |
| 22 | 40.4 | 8.4 | 41118 | 2 | AC090895 | AC090895 Homo sapi |
| 23 | 40.2 | 8.4 | 106559 | 2 | AC010295 | AC010295 Homo sapi |
| 24 | 40 | 8.3 | 7448 | 3 | AF465305 | AF465305 Dictyoste |
| 25 | 40 | 8.3 | 18446 | 8 | AB026660 | AB026660 Arabidops |
| 26 | 40 | 8.3 | 19132 | 2 | ATH131517 | ATH131517 Arabidops |
| 27 | 40 | 8.3 | 157672 | 8 | AC100742 | AC100742 Mus muscu |
| 28 | 40 | 8.3 | 178493 | 2 | AL354981 | AL354981 Homo sapi |
| 29 | 39.8 | 8.3 | 11321 | 1 | AE006407 | AE006407 Lactococc |
| 30 | 39.8 | 8.3 | 11541 | 1 | AE007813 | AE007813 Clostridi |
| 31 | 39.8 | 8.3 | 74589 | 8 | AB010068 | AB010068 Arabidops |
| 32 | 39.4 | 8.2 | 12668 | 3 | AF079445 | AF079445 Dictyoste |
| 33 | 39.4 | 8.2 | 58753 | 3 | AF125459 | AF125459 Caenorhab |
| 34 | 39.4 | 8.2 | 164051 | 9 | AC106892 | AC106892 Homo sapi |
| 35 | 39.2 | 8.1 | 145938 | 2 | AC128165 | AC128165 Rattus no |
| 36 | 39 | 8.1 | 4183 | 3 | AF474377 | AF474377 Dictyoste |
| 37 | 39 | 8.1 | 180854 | 9 | AC107960 | AC107960 Homo sapi |
| 38 | 38.8 | 8.1 | 197092 | 2 | AC036155 | AC036155 Homo sapi |
| 39 | 38.8 | 8.1 | 6294 | 6 | AX344308 | AX344308 Sequence |
| 40 | 38.8 | 8.1 | 9997 | 6 | AX281373 | AX281373 Sequence |
| 41 | 38.8 | 8.1 | 189235 | 2 | AC113717 | AC113717 Rattus no |
| 42 | 38.8 | 8.1 | 195411 | 2 | AC092508 | AC092508 Mus muscu |
| 43 | 38.6 | 8.0 | 85421 | 8 | AB007651 | AB007651 Arabidops |
| 44 | 38.6 | 8.0 | 149715 | 2 | AC117165 | AC117165 Rattus no |
| 45 | 38.6 | 8.0 | 149715 | 2 | AC117165 | AC117165 Rattus no |

ALIGNMENTS

```
RESULT 1
AX278241 481 bp
LOCUS AX278241
DEFINITION Sequence 1 from Patent WO0177176.
ACCESSION AX278241
VERSION AX278241.1 GI:16605287
KEYWORDS
SOURCE
ORGANISM Escherichia coli.
Bacterioph: Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1
AUTHORS Levy,S.B. and McMurry,L.M.
TITLE Novel b1r molecules affecting antibiotic susceptibility
JOURNAL Patent: WO 0177176-A 1 18-OCT-2001;
```

TRUSTERS OF TUFTS COLLEGE (US)
 FEATURES Location/Qualifiers
 source 1. 481
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 CDS 239..364
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD10539.1"
 /db_xref="GI:16605288"
 /translation="MNRIELTGMIVLVVSYLLGVASHIDNPPEOSASVQHK"
 BASE COUNT 139 a 95 c 100 g 147 t
 ORIGIN
 Query Match 100.0%; Score 481; DB 6; Length 481;
 Best Local Similarity 100.0%; Pred. No. 2e-115;
 Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCCTCTCATTTAGATGATGCTATTGACATATCCTTTACCCAGCTCAAGCTTT 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 TGCCTCTCATTTAGATGATGCTATTGACATATCCTTTACCCAGCTCAAGCTTT 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 AATAACCTGCCAGCAATAAGGAGATGTTGTTAACTTAAGTCAAAAAAATAGCGAATTTTC 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 AATAACCTGCCAGCAATAAGGAGATGTTGTTAACTTAAGTCAAAAAAATAGCGAATTTTC 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 CAACGACAAAGCTAAATATCGCAAAAACCTCAGTAAAAATCTCTGAGCTATTAT 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 CAACGACAAAGCTAAATATCGCAAAAACCTCAGTAAAAATCTCTGAGCTATTAT 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 GCTAAGTAACTTACCCCTGAGTAAATGATCATCAAGAGATGGGCTGAAT 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 181 GCTAAGTAACTTACCCCTGAGTAAATGATCATCAAGAGATGGGCTGAAT 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 GAATCGTCTTATTTGAATTAACAGGTGGATCGTCTGTCGTTTCAGTATCTTCTGG 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 GAATCGTCTTATTTGAATTAACAGGTGGATCGTCTGTCGTTTCAGTATCTTCTGG 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 CGTGGGAGTGCACATTGACAACTATACGCCACTGTGAACAGAGTGGTGTACAAACAA 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 301 CGTGGGAGTGCACATTGACAACTATACGCCACTGTGAACAGAGTGGTGTACAAACAA 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 GTAAGCTGTCACATTGAGAGACATGCGCCGTCGCGGTCATGTTTCACTTGTGC 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 361 GTAAGCTGTCACATTGAGAGACATGCGCCGTCGCGGTCATGTTTCACTTGTGC 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 GATATTAACCAAGAAATTTATATCTTGTTCATGATGTTGTTGATTTGCAGGATAGTA 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 421 GATATTAACCAAGAAATTTATATCTTGTTCATGATGTTGTTGATTTGCAGGATAGTA 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 481 A 481
 ||
 DB 481 A 481
 ||

RESULT 2
 AE000258 13172 bp DNA linear BCF 01-DEC-2000
 LOCUS Escherichia coli K12 MG1655 section 148 of 400 of the complete genome.
 DEFINITION
 accession AE000258.1 GI:2367121
 version AE000258.1
 keywords
 source Escherichia coli K12.
 organism Escherichia coli K12.
 organism Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 reference 1 (bases 1 to 13172)
 authors Blatner, F.R., Plunkett, G., III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Colado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., and Shao, Y.

TITLE The complete genome sequence of Escherichia coli K-12
 JOURNAL Science 277 (5331), 1453-1474 (1997)
 MEDLINE 97426617
 PUBMED 9278503
 REFERENCE 2 (bases 1 to 13172)
 AUTHORS Blatner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
 REFERENCE 3 (bases 1 to 13172)
 AUTHORS Blatner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
 REFERENCE 4 (bases 1 to 13172)
 AUTHORS Plunkett, G., III.
 TITLE Direct Submission
 JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 COMMENT On Sep 9, 1997 this sequence version replaced gi:1787907.
 This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: markborov@ga.tech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES
 source Location/Qualifiers
 1. 13172
 /organism="Escherichia coli K12"
 /strain="K12"
 /sub_strain="MG1655"
 /db_xref="taxon:83333"
 76..1668
 /gene="malX"
 /note="b1621"
 76..1668
 /gene="malX"
 /function="Transport; Transport of small molecules: carbohydrates, organic acids, alcohols"
 /note="o530; 99 pct identical to PROA-ECOLI SW: P19642; CG Site No. 32254; phosphotransferase system enzyme II"
 /codon_start=1
 /transl_table=11
 /product="PTS system, maltose and glucose-specific II ABC"
 /protein_id="AAC74693.1"
 /db_xref="GI:1787908"
 /translation="MTAKTAPVTLMEFPOOLGKFTMLPVALLSPCGIMLGISLSS HVTILIPVGNPVLQALFTWMSKIGSFASFLPVMFCIALPLGARLNKGVAAAFAG


```
IGYAVMNLAVNFWLTKGILPTTDAVLKANNIISLIGIOSIDRGILGAVTACITVM
LHERPHNIRLPDALAFEGGTRRVPIITSSVMKLVLYPLWPIRAMKISGSHINS
AGDEGPMLEIGTERILLPGLHILVALIRFDAGTQVCGQIVSGALITRQALSC
PTTHGFSSESATFELSQGMKPAFLGGLPGALAMYCARENRHKIKGLISGLICVY
GTTTEPLEFLEFLVAPVLYVIALITGLGFTVMVSLGATIGNTDENIIDFVYGLIH
LSTKVMVPVVAIWEVYVYFRAITRENLKTRGRSEVSSIEKAVAGAPGSGY
NVPALTEALGGADNIVSLDNCITRRLRSVKMSLYNVQALKNRRAIGVQLNQHNLY
VTGPVOSVKDEMAGLMHTVQA"
promoter
/gene="maly"
/ncr="factor Sigma54; predicted +1 start at 1638961"
1678..2850
/ncr="maly"
/ncr="b1622"
1678..2850
/gene="maly"
/ncr="factor Sigma2.6.1.-"
/function="enzyme; Degradation of small molecules: Carbon
compounds"
/ncr="0390; 100 pct identical to Maly_ECOLI SW: P23256;
CG Site No. 32257"
/codon_start=1
/transl_table=11
/product="enzyme that may degrade or block biosynthesis of
endogenous mal inducer, probably aminotransferase"
/protein_id="AAC74694.1"
/db_xref="GI:1787909"
/translation="MEDSKVDRHGTCTQMDVYARFGTADLLPTTISDMDFATAP
CIEALNORLMHGVGYSRKNDEFLLAIHMFSTOHTAIDSOYVYVPSYIVYASE
LIRWSETEGEVVIHTPAYDAFYKAIEGNOTRTVMVLAIDSGVMDGKKEAVLAK
PECKIMLCSQNPRTGKWTGDELMADICRHRGVYVLSDEIHMDMGWDEPHIAMS
NYARGDMLITSGSKSPNKPATGAYGIENSSSDAYLSAKRGDGLSSPEVIALTA
HIAAYQCGAPWLDALRIYKMDLITADKMAAFELMNOIPSTYLMADIRPLNID
DNALOKALIEDGKVAIMPYTYGEEGRFRLNACGPKSKLEKGVAGLINALRAVR"
2887..2916
/ncr="factor sigma70; predicted +1 start at 1700226"
2954..3955
/gene="add"
/ncr="b1623"
2954..3955
/gene="add"
/ncr="factor Sigma3.5.4.4"
/function="enzyme; Salvage of nucleosides and nucleotides"
/ncr="0333; This 333 aa ORF is 99 pct identical (1 gap)
to333 residues 332 aa protein ABD_ECOLI; CG Site No. 1042
SW: P22333"
/codon_start=1
/transl_table=11
/product="adenosine deaminase"
/protein_id="AAC74695.1"
/db_xref="GI:1787910"
/translation="MIDTTLPLTDIHRHLNDNIRPOTIIELGROYNISLPAQSLTLL
PMAOVIANPDLVSELTAKDMGVKVLASDACRRAVENIEDAARGGLYVLRSPG
PYMAHOLPAGVAVAVIDVREGRCTFVQAKLIGIMRTGEAGACCOELAEALNR
DOITLADLAGDELGFPGSLFLSHFNARARAGWHIYVAGEAGPESIMOAIEELGAR
IGHGVKAIEDRLMDPIAEQOIGIESCLTNSIOTSGYVELAHPRLKTELEHRSIN
TDDPGVQGVDIHETVAVAPAGLSREBQIRQOINGLEMAFLSAEKKRLRKVAAK"
complement(3989..5068)
/gene="b1624"
complement(3989..5068)
/gene="b1624"
/function="orf; Unknown"
/ncr="f359; residues 74-259 are 45 pct identical to
aa62-244 from hypothetical protein YNH_ECOLI SW: P46853
(345 aa)"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC74696.1"
/db_xref="GI:1787911"
/translation="MSNSHTSMRGNDSNIRVGLIGYVASKTFHAPLIAGTPOGE
LAVISSDETKVADMPITVYVSEPKHLENDILIDLPYNDHIFPLAKALDEAGK
HYVVDKPFVTLSQARELDALAKSLGRLVSFHNRRMWDPLTLGLLAEGVLEAVY
```

```
PSHRRPREPOVRDRBREGGSGIWDIAPHILDOATLEGLPVSMTVDLAQLRPG
AOSTDYFNALISYPOVRVILHCTMTAAESARYTHGSGSYVKGLDPOERLKNGE
RUPQEDWCVLDRVETLREGEERVELLTVPGVYPAAYAIRDALNGDGENPVAS
QAIVQWELIELIESAKHRATCILA"
repeat_region
/ncr="REP (repetitive extragenic palindromic) element;
contains 1 REP sequence"
4124..4158
promoter
/complement(5165..5193)
/ncr="factor Sigma70; predicted +1 start at 1702461"
5498..5526
promoter
/ncr="factor Sigma70; predicted +1 start at 1702836"
5556..5582
promoter
/ncr="factor Sigma70; predicted +1 start at 1702892"
5670..5885
gene
/gene="b1625"
5670..5885
/gene="b1625"
/function="orf; Unknown"
/ncr="071; This 71 aa ORF is 42 pct identical (4 gaps)
to71 residues of an approx. 72 aa protein YMOA_YEREN SW:
p27720"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC74697.1"
/db_xref="GI:1787912"
/translation="MTVDYLLKFRKISSLESLEKLYDHLNNTLTDDDELIMYRAAD
HRAELVSGGRLEFDGQVSVWYHQ"
5755..5783
/gene="b1625"
/ncr="factor Sigma70; predicted +1 start at 1703093"
5904..5933
promoter
/ncr="factor Sigma70; predicted +1 start at 1703243"
5947..6411
gene
/gene="b1626"
5947..6411
/gene="b1626"
/function="orf; Unknown"
/ncr="0134; This 134 aa ORF is 27 pct identical (5 gaps)
to134 residues of an approx. 135 aa protein YMOA_YEREN SW:
p27720"
Query Match 97.5%; Score 469; DB 1; Length 13172;
Best Local Similarity 99.8%; Pred. No. 2.5e-112;
Matches 480; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 TGCCTCTCATGTAAGTATGATGCTATTTGACATATCCTTAACCAAGCTCAAGCTT 60
|||||
DB 5035 TGCCTCTCATGTAAGTATGATGCTATTTGACATATCCTTAACCAAGCTCAAGCTT 5094
QY 61 AATTAACCTGCACCAATTAAGGATGTTGTTTAACTTAAGCAAAAATATGCAATTTTC 120
|||||
DB 5095 AATTAACCTGCACCAATTAAGGATGTTGTTTAACTTAAGCAAAAATATGCAATTTTC 5154
QY 121 CAACGAACAACAACTAATATCGCAAAAACCTGAGTAAATCTTGCGAGCTATTATT 180
|||||
DB 5155 CAACG-ACAAACCTAATATATGCAAAAACCTGAGTAAATCTTGCGAGCTATTATT 5213
QY 181 GCTAAGTAACTTTACCCCTGAGTTAATGATCAATCAAGAGATGTGGCTGTAAAT 240
|||||
DB 5214 GCTAAGTAACTTTACCCCTGAGTTAATGATCAATCAAGAGATGTGGCTGTAAAT 5273
QY 241 GAATCGCTTATGAAATTAACAGTTGATCGTCTTGCTGCTGCTGCTGCTGCTG 300
|||||
DB 5274 GAATCGCTTATGAAATTAACAGTTGATCGTCTTGCTGCTGCTGCTGCTGCTG 5333
QY 301 CGTGGGAGTACATTTGACAACTATGACGACCTGAAACAGAGTCTTCGTACACACAA 360
|||||
DB 5334 CGTGGGAGTACATTTGACAACTATGACGACCTGAAACAGAGTCTTCGTACACACAA 5393
QY 361 GTAAAGCTTGTGACTTGTGAGCGACATGCTCCCGTCCGGGTGCATGTTTTCATCTG 420
|||||
DB 5394 GTAAAGCTTGTGACTTGTGAGCGACATGCTCCCGTCCGGGTGCATGTTTTCATCTG 5453
QY 421 GATATTAAACCAAGATTTTATATCTTGTTCGATGTTGTTGGATTTGTTCAGGATTA 480
```

|||||

DB 5454 GATTATTAAACCGAATTATTATCTGTGATGTGTGTGATGTGCAGGATAGTA 5513

QY 481 A 481

DB 5514 A 5514

RESULT 3

D90806 15332 bp DNA linear BCT 29-MAY-1997

LOCUS D90806 E.coli genomic DNA, Kohara clone #315(36.6-36.9 min.).

DEFINITION D90806 AB001340

ACCESSION D90806.1 GI:1742679

VERSION D90806.1

KEYWORDS Complete and shotgun sequencing; add: hsdh, hsdh; malI; malX; malY; ntl.

SOURCE Escherichia coli (strain:K12) DNA, clone_lib:Kohara lambda miniset library clone:Kohara clone #315.

ORGANISM Escherichia coli

REFERENCE Escherichia coli

AUTHORS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

1 (sites)

Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.

The systematic sequencing of the Escherichia coli genome in Japan Unpublished 2 (bases 1 to 15332)

Mori,H

Direct Submission

Submitted (14-DEC-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5 Takayama, Ikoma, Nara 630-01, Japan

(E-mail:hmori@gtc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)

Collaboration Information:

Project:

The Japan E.coli genome DNA sequencing project

Group:

The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)

Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.

Headed by:

Name: Takashi Horiiuchi

Address: National Institute of Basic Biology, Okazaki, 444, Japan

E-mail: kishori@nibb.ac.jp

Information operator:

Name: Hirotsada Mori

Address: NARA Institute of Science and Technology, Ikoma, 630-01, Japan

E-mail: hmori@gtc.aist-nara.ac.jp

URL:

The Japan E. coli genome database

http:bsw3.aist-nara.ac.jp

Location/Qualifiers

1. .15332

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/map="36.6 min"

/clone="Kohara clone #315"

/clone_lib="Kohara lambda miniset library"

CDS

/note="Nucleotide position 1699836-1715167 from the initiation site of Thra (0 min.). This clone is from Kohara lambda miniset library"

complement(<1.48)

/note="ORF_ID:0314#4

similar to [Swissprot Accession Number P34000]"

/codon_start=1

/transl_table=1

/product="Potential acraB operon repressor."

/protein_id="BA15376.1"

/db_xref="GI:1742680"

/translation="TMDNQTEAQPTRRI"

complement(269..1036)

/gene="hdhA, hsdh"

complement(269..1036)

/gene="hdhA, hsdh"

/note="ORF_ID:0314#5

similar to [Swissprot Accession Number P25529]"

/codon_start=1

/transl_table=1

/product="7-a-hydroxysteroid dehydrogenase (EC 1.1.1.159)

(7.alpha.-HSDH)"

/protein_id="BA15377.1"

/db_xref="GI:1742681"

/translation="MFNSDNLRLDKCAIITGACAGIGKEIATPATGASVYSDIN

ADANHYVDEIQGLGQAFACRCDITSOEISALADPAISLGRVDILVNNAGGGRK

PEQPMDFERRAYELNVESEFHLISQIVAEKNGCGVILITISMAENNNIMTSYA

SSKKAASHLVNNAFMDGKRNIRYNGIAPGAILDLDAKSVTPPEIQMKLHPHRL

COPDINAAALFICSPASWSGQILVSGGVELN"

complement(1148..2176)

/gene="malI"

complement(1148..2176)

/note="ORF_ID:0314#6

similar to [Swissprot Accession Number P18811]"

/codon_start=1

/transl_table=1

/product="Maltose regulon regulatory protein MalI."

/protein_id="BA15378.1"

/db_xref="GI:1742682"

/translation="MTAKKITIHVDVALAGSVSYSLVLSGKRGISTATGEVNNA

IEELGFVNNRQASALRGQSGVIGLIYRDISAPFYAELTGLTALRAQGRMYLLDG

GRDEQLAORFSLNLNGVDGVILGANGSSDDLPRAAEKATVITASRSGYDDPD

TVRPDPMQAOLITHEILIRNGHRIAWLGGSSSLTRAERVGTCATLLKGLPFPHSD

WVLECTSSQKQAAEAITALLRHNPISAVVQYNETIAGWFGLLKGRSGESGVIR

YEEQGVSLAQTADATPTLDDIPVTWASTPAREIGITLADRMOKITHEETHSRNLLI

PARLIAAK"

2351..3943

/gene="malX"

2351..3943

/gene="malX"

/note="ORF_ID:0314#7

similar to [PIR Accession Number B42477]"

/codon_start=1

/transl_table=1

/product="Phosphotransferase system enzyme II (EC 2.7.1.69) MalX"

/protein_id="BA15379.1"

/db_xref="GI:1742683"

/translation="MTAKTAPVPTLMEFFOOLGKTFMLPVALLSFCGIMIGSSLS

HDVITLIVGPNPVOAIFETWMSKIGSAFSLPMPFCIAPJGLARENKGVAFAG

ICYAVMNIAVNNFWLNNKGIITPTDAVAVKANNISIIIGISIDIGIGAGVAGIIVWM

LHERHNIRLDPALAFEGGTRFVPIISSLVAGLVGLVPLVPYPMFAMGISLGHINS

AGDFGPMLEGTGERLLPFGLLHIIVALIRFTDAGCTQEGGVSGALTFQALQSC

PTTGFSESATRELSQGMKPAFELGGLPGALAMYGHCAPENRHKIKGLISGLACV

GGTEPLELEFLFVAPVLYVYIHALLTGLGFTVMVSVLYTIGNTDGIIDFVFGILHG

ISTKRWMPVVAIWEVYVYVFERAIPFNKTPGGRSEVASSIEKAVACAPGSGY

NVPALIEALGADNTVSDNCTIRLRISVKKMSLVNQAALDNRAIGVQVQNGHMLVY

VIGPVQVSXKDMAGLMTVQA"

3953..5125

/gene="malY"

3953..5125

/gene="malY"

gene
CDS

```

/note="ORF_ID:0314#8
similar to [PIR Accession Number C42477]"
/codon_start=1
/transl_table=1
/product="Malt protein"
/protein_id="BA15380.1"
/db_xref="GI:1742684"
/translation="MFDESKYVDRHRTWCOTQMDYVADRFETADLLPTISMDPFTAP
CIIFALNORLMHGVGYSRMKNDPEFLAIAHPESTOHTAIDISOTVYGPSEVITWSE
LIRWSEGESEVYHTPAYDAFYKAIBENQNTNMPALEKADQDFCDMGKLEAVLAK
PECKIMLCSPONTGKVTCDDELEIMADLERGVRVTSDEIHMDYWGQPHIPLWS
NVAAGDAPMLDRLRYIKDNLNTYADKNMAAPELMQIPOSTYLAIMDLRPLND
DNALOKALIEQEKVAIMPGYTYGEGRGFVRLNMGCPRSKLEKGYAGLINRAVR"
5229..6230
/gene="add"
/gene="add"
5229..6230
/gene="add"
/note="ORF_ID:0314#9
similar to [SwissProt Accession Number P22333]"
/codon_start=1
/transl_table=1
/product="Adenosine deaminase (EC 3.5.4.4) (Adenosine
aminohydrolase)."
/protein_id="BA15381.1"
/db_xref="GI:1742685"
/translation="MIDTFLPLDHRHLDGNIRPORTLEIROYNTSLPASOSETLI
PHVOYIANEPDVSFLTKLDMGVYKASLDCRRAVENIDDAARHGHIYELRSPG
YMAAHQPLVAGVYEAVIDGVREGKRTFVQAKLIGISRFEGEACQOLEAFLAHR
DOITALDAGDELGPGLSLFSEHNRADAGMHTVHAGEAPEGSIWQAIREFLAER
IGHVVKALIEDRALMDFLAEQOIGIESCLTISQSTVLAELAHPLKPTLEHIGIRASIN
TDDPGVOGVDIRHEHYVAAPAGLSREQIRQAOJNGLEMAFLSAEKRALREKVAAK"
complement(6264..7343)
/note="ORF_ID:0314#10
similar to [SwissProt Accession Number P37168]"
/codon_start=1
/transl_table=1
/product="Virulence factor Mv1M."
/protein_id="BA15382.1"
/db_xref="GI:1742686"
/translation="MSNSHRTSMRGNDSDNIRVGLIGYGYATFHAAPLIAGTPGOE
LAVISSDETQVKAQWPTVYSEPKHLFNDPNIDLIYIPNDTHPELAKAALEAGK
HVVVDKPTVTSOARELDALAKSLGRVLSYEHNRMSDPLTLGLAEVGLGVA
FESHDRPPOVRDRMRROGGSGGIWYDLAHLDDOITLFGLPVSMYUQLALRPG
AOSTDYFALISTYQORYILGHTMAAESRYIVHSGRSYVKGIDPOERLNGE
RLPOEDMGIDMRDGLVTRVEGEVEETLLTPGNYPAYVAIRDALNGDENPVAS
QAIVMELEIGIESAKHRAFLCLA"
9344..9922
/note="ORF_ID:0316#4
similar to [SwissProt Accession Number P00211]"
/codon_start=1
/transl_table=1
/product="Peredoxin II."
/protein_id="BA15383.1"
/db_xref="GI:1742687"
/translation="MNAWIVAAVSLGLAFGALIGYASRPVADDPVYEKIDELI
POSOGGCGTGPCRPAISCGNEKINRCAPGEAVWLKIAELINVEPOPULDEAOE
ITPARVAVIDENNCIGCTKCIOACPDVAIVGAIARHMTVMSDLCTGNCICVDCPTH
CISLPVAIEPDPDSKWDINTIPRIIPEVHA"
9915..12137
/note="ORF_ID:0316#5
similar to [SwissProt Accession Number P14922]"
/codon_start=1
/transl_table=1

```

Query Match 97.5%; Score 469; DB 1; Length 15332;
 Best Local Similarity 99.8%; Pred. No. 2,4e-112;
 Matches 480; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 TGCCTCATTTGAGATGATGATGATTTGACATATCCTTTACCAAGCTCAAGAGTTT 60
 DB 7310 TGCCTCATTTGAGATGATGATGATTTGACATATCCTTTACCAAGCTCAAGAGTTT 7369

```

OY 61 AATAACCTCCAGCATATAAGGATGTGTTTAACTTAAGTCAAAAAAATAGCGAATTTTC 120
DB 7370 AATAACCTCCAGCATATAAGGATGTGTTTAACTTAAGTCAAAAAAATAGCGAATTTTC 7429
OY 121 CAACGACAAAAGCTAAATATATGCAAAAAACCTCAGTAAAAATCTTGCTGAGACTATTAT 180
DB 7430 CAACG-ACAAAAGCTAAATATATGCAAAAAACCTCAGTAAAAATCTTGCTGAGACTATTAT 7488
OY 181 GCTAAGTAACATTTACCCCTCAAGTAAATGATCAATCAACAGAGATGTGGCTGTAA 240
DB 7489 GCTAAGTAACATTTACCCCTCAAGTAAATGATCAATCAACAGAGATGTGGCTGTAA 7548
OY 241 GAATGCTCTTATGATTAACAGGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 7549 GAATGCTCTTATGATTAACAGGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7608
OY 301 CGTGGCGAGTCACATTTGACACATATACAGCCACCTGACAGAGTGTGCTGCTGCTGCTGCTGCT 360
DB 7609 CGTGGCGAGTCACATTTGACACATATACAGCCACCTGACAGAGTGTGCTGCTGCTGCTGCTGCT 7668
OY 361 GTAAGCTCTGCACTTGTGGAGCGACATGCTGCCGCTGCCGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 7669 GTAAGCTCTGCACTTGTGGAGCGACATGCTGCCGCTGCCGCTGCTGCTGCTGCTGCTGCTGCTG 7728
OY 421 GATATTAAACCGAATTTATATCTGTTGATGTTGTTGATGTTGTTGATGTTGTTGATGTTGTTG 480
DB 7729 GATATTAAACCGAATTTATATCTGTTGATGTTGTTGATGTTGTTGATGTTGTTGATGTTGTTG 7788
OY 481 A 481
DB 7789 A 7789

```

RESULT 4
 D90808 15601 bp DNA linear BCT 29-MAY-1997
 LOCUS E.coli genomic DNA, Kohara clone #317(36.6-36.9 mln.).
 DEFINITION D90808 AB001340
 ACCESSION D90808.1 GI:1742711
 VERSION Complete and shotgun sequencing; add; hbha, hsdh; malI; malX; malY;
 ntl.
 SOURCE Escherichia coli (strain:K12) DNA, clone_1lb:Kohara lambda miniset
 library clone:Kohara clone #317.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE
 1 (sites) Alba,H., Baba,T., Fujita,K., Hayashi,K., Inada,T., Isono,K.,
 Itoh,T., Kasai,H., Kashimoto,K., Kimura,S., Kitakawa,M.,
 Kitagawa,M., Makino,K., Miki,T., Mizobuchi,K., Mori,H., Mori,T.,
 Motomura,K., Nakade,S., Nakamura,Y., Nashimoto,H., Nishio,Y.,
 Oshima,T., Saito,N., Sempel,G., Seki,Y., Sivasubram,S.,
 Tagami,H., Takeda,J., Takemoto,K., Takeuchi,Y., Wada,C.,
 Yamamoto,Y. and Horiiuchi,T.
 A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28,040,1 min region on the linkage map
 DNA Res. 3 (6), 363-377 (1996)

TITLE
 JOURNAL MEDLINE
 REFERENCE
 2 (sites) Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T.,
 Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K.,
 Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M.,
 Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H.,
 Motomura,K., Nakamura,Y., Nashimoto,H., Nishio,Y., Oshima,T.,
 Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
 Yamamoto,Y. and Yano,M.
 The systematic sequencing of the Escherichia coli genome in Japan
 Unpublished
 3 (bases 1 to 15601)

TITLE
 JOURNAL
 REFERENCE
 3 (bases 1 to 15601)
 Direct Submission
 Submitted (14-DEC-1996) Hirotada Mori, NARA Institute of Science
 and Technology, Res. & Edu. Center for Genetic Info.; 8916-5

COMMENT

Takayama, Ikoma, Nara 630-01, Japan
(E-mail:hmori@etc.aist-nara.ac.jp, Tel:81-7437-2-5660,
Fax:81-7437-2-5669)
Collaboration Information:

Project:

The Japan E.coli genome DNA sequencing project

Group:

The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)

Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A.,
Horluchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S.,
Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S.,
Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K.,
Masuda, S., Miki, T., Mizoduchi, K., Mori, H., Motomura, K.,
Nakamura, Y., Nashimoto, H., Nishio, Y., Oshima, T., Saito, N.,
Sampel, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,
Yamamoto, Y. and Yano, M.

Headed by:

Name: Takashi Horluchi

Address: National Institute of Basic Biology, Okazaki, 444, Japan

E-mail: kishori@nibb.ac.jp

Information operator:

Name: Hirotsada Mori

Address: NARA Institute of Science and Technology,

Ikoma, 630-01, Japan

E-mail: hmori@etc.aist-nara.ac.jp

URL:

The Japan E. coli genome database

http://sw3.aist-nara.ac.jp.

Location/Qualifiers

1. 15601

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/map="36.6 min"

/clone="Kohara clone #317"

/clone_id="Kohara lambda miniset library"

/note="Nucleotide position 1699735-1715335 from the

initiation site of Phra (0 min.).-This clone is from

Kohara lambda miniset library"

complement(1. 149)

/note="ORF_ID:0314#4

similar to [SwissProt Accession Number P34000]"

/codon_start=1

/transl_table=11

/product="potential acraB operon repressor."

/protein_id="BA15406.1"

/db_xref="GI:1742712"

/translation="MMDMOTEAQPTRIILNAREIFSENGFHSASKAICSAIS

PGTLY"

complement(370. 1137)

/gene="hsha, hsdh"

complement(370. 1137)

/gene="hsha, hsdh"

/note="ORF_ID:0314#5

similar to [SwissProt Accession Number P25529]"

/codon_start=1

/transl_table=11

/product="7-a-hydroxysteroid dehydrogenase (EC 1.1.1.159)

(7-alpha-HSDH) "

/protein_id="BA15407.1"

/db_xref="GI:1742713"

/translation="MNSDNLRLDGKCAIITGAGAGIGKEIAITFATAGASVVDIN

ADANHVDEIQQLGQAFACRCDITSEQLSALADFAISKLGVDILVNNAGGGPK

PEMDADFERAYENVSFPHLSQIVPEKNGGVILITSSAAENKINNTSYA

SSKAASHLVNNMAFDGKRNKNGTAPGAILDALKSVITPECKMLQHTIRRL

GQPDIAAAALFLCSPPASVSGQILTVSGGVDPLN"

complement(1249. 2277)

/gene="mali"

complement(1249. 2277)

/gene="mali"

/note="ORF_ID:0314#6

similar to [SwissProt Accession Number P18811]"

/codon_start=1

/transl_table=11

/product="Maltose regulon regulatory protein Mali."

/protein_id="BA15408.1"

/db_xref="GI:1742714"

/translation="MAFAKKITIHVALAGVSVSVLSLGGKRGISTATGEVYAA

IEELGIVRNROASALRGOSGVGLITRDLSPAFVAFETAPGLETALAEAGRWFLHG

GKDRGLAAREFILLNOGVADVAGVIAAGSSDDLRMAEKAIPVIEPARASYLDVD

TVAPDMQAQILTEHLIRNGHORINMLGCOSSSTRARNGGCATLLKFLPFSD

WVECHSSRQAALVATLRNPTISAVVCNETHAKAMGILKAGCGSEGVDR

YFEQVSLAATDPTTDDIPVWASPARELITTLADRMQKITHEHTRNLII

PARLIAX"

2452. 4044

/gene="maliX"

2452. 4044

/gene="maliX"

/note="ORF_ID:0314#7

similar to [PIR Accession Number B42477]"

/codon_start=1

/transl_table=11

/product="phosphotransferase system enzyme II (EC

2.7.1.69) MaliX"

/protein_id="BA15409.1"

/db_xref="GI:1742715"

/translation="MTAKTAPKVTLMFEFQQLGKPEMLPVALLSFCCIMLGISLSS

HDVITLIPVGNFVLOAIFWMSKIGSFASFPLVMECAIPGLARENKGYAAEAFG

IGYAVNNIAVNLNFKGILPTDAVLANNNISILGOSIDTGLGAVGIYTM

LHERFNIRLPDALARFGGRFVPIISLVMGVGLVPLMPTIFAMGISGGMINS

AGDFGMLRGTGRLRLPRGLHIIIVALLIRPTDAGTGVCCGTGALITFOALSC

PTTHGFSESATRFSLGKMPALFGLGALALYHCAREBNRKHIGLISLILACY

GGTTEPLEFLFVAPVLYTHALGLGTWVSLGVITGNDGNIIDVVFGLIG

LSTRKVMVPVVAIMEVYVIVIRFATIRNLKTPGRDSEVASSIERAAVAGPSKY

NVPAILEALGADNIVSLDNCITRLRLSVKMSLVVVALKDNRAIGVVOLOHNLQV

VIGPOVSVKDEAGLMHTVOA"

4054. 5226

/gene="maliY"

4054. 5226

/gene="maliY"

/note="ORF_ID:0314#8

similar to [PIR Accession Number C42477]"

/codon_start=1

/transl_table=11

/product="MaliY protein"

/protein_id="BA15410.1"

/db_xref="GI:1742716"

/translation="MDEFSKVVVRHGTQWMDVADREGTADLLEPTISDMDFATAP

CIEALNORLMHGVGYRWNKDEFLAIAHMFSTOYHAIIDQTVVYGSVLYVSE

LIRWSEFGVVIHPAYAPAFKAIEGNQRTVPALEAIOAGVFCDMGKLEAVLAK

PECKIMLCSPPONPTGVKWCDELEIMADICRHGVRVTSDEIHMVMVGEOPHIPS

NVARGDALITSGSKFNIPALTGAIGIENSSRDPAIYSAIKRGGLSSPSYLAITA

HIAVYQGAIPWDLALRYLKDNLTYLADKNNAFPELNMQIPOSTYLAWDLPLNID

DNALOKALIEQEVYALMPGYTGBEGGFVRLNAGCPRSKLEKVGALINAIARV"

5330. 6331

/gene="add"

5330. 6331

/gene="add"

/note="ORF_ID:0314#9

similar to [SwissProt Accession Number P22333]"

/codon_start=1

/transl_table=11

/product="Adenosine deaminase (EC 3.5.4.4) (Adenosine

aminohydrolase) "

/protein_id="BA15411.1"

/db_xref="GI:1742717"

/translation="MIDTFLPDIHRHLDGNIRPQITELRGQYNISLPAQSLLETLI

PHYVAINEDDLVSEFLKLMGVYVSLADCAVRPAEENEDAAHGLHAYELRFSPG

YMAAHOGLPAGVYEAVIDVREGCRFFGVQAVLIGMSTRFEGAACQOQLAEFLAR

DOITADLADDELGFPSGLSLSHNRARDGWHYVAGCAAGPESTWQIIRLGAR

IGHGVKAIEDRALMDFLAEQIGISLTSINQISTYAEIAAHPKRTFHEHGRASIN

TDPCGVGVDIIRHYTYVAPAPAGISRQIRQAOINGLEMALSAEERALARVAAK"

complement(6365. 7444)

/note="ORF_ID:0314#10

similar to [SwissProt Accession Number P37168]"

```

/codon_start=1
/transl_table=11
/product="Virulence factor mw1m."
/protein_id="BA15412.1"
/db_xref="GI:1742718"
/translation="MSNSHHTSMRGNDMSDNIRVGLIGYVASKTFHAPLIAGTPEOE
LAVISSDETKYKADWPVTYVSEPKREHNDNDIVIPFNDFHPLAKALEGK
HYVVDKPTVTYLSOARELDALAKSIGRYSVFNHNRWSDPFTNDGILAEVLEKAY
FESHDFRFPQVDRKDRREGGSGGWDLAHLDDAITLGLPYSMTVDLAQLKPG
AOSTDYFHALISYPOKRVTLHGTMLAASARYIVHSGSYVKYGLDQERLNKE
RLPOEDWGYMDRGVLTREVEGERVEETLLTVPGNYPAYAAIRDALNDGENPVPAS
OAIQVMELEICIESAKHRATLCIA"
9445.10023
/contig="ORF_ID:0316#4
similar to [SwissProt Accession Number P00211]"
/codon_start=1

Query Match      97.5%; Score 469; DB 1; Length 15601;
Best Local Similarity 99.8%; Pred. No. 2.4e-112;
Matches 480; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 TGCCTCATTTGAATGATGCGCTATTGTGACACTATCTTTACCCAGCTCAACAGTTT 60
    |||||||
DB 7411 TGCCTCATTTGAATGATGCGCTATTGTGACACTATCTTTACCCAGCTCAACAGTTT 7470

QY 61 AATAACCTGCAGCAATAAGGAGATGTTGTTAACTTAAGTCAAAAAAATAGCGAATTTTC 120
    |||||||
DB 7471 AATAACCTGCAGCAATAAGGAGATGTTGTTAACTTAAGTCAAAAAAATAGCGAATTTTC 7530

QY 121 CAACGAACAAGCTAAATATCGCAAAACCTGAGTAAATATCTGCTGAGCTATTATT 180
    |||||
DB 7531 CAACG-ACAAAGCTAAATATCGCAAAACCTGAGTAAATATCTGCTGAGCTATTATT 7589

QY 181 GCTAAGTAACTTACCCCTGAGTAAATGATGATCAATCAAGAGATGTGGCTGTAA 240
    |||||||
DB 7590 GCTAAGTAACTTACCCCTGAGTAAATGATGATCAATCAAGAGATGTGGCTGTAA 7649

QY 241 GAATCGCTTATGAAATTAACAGTGGATGCTGCTTGTGCTTCACTATTCTTCTG 300
    |||||||
DB 7650 GAATCGCTTATGAAATTAACAGTGGATGCTGCTTGTGCTTCACTATTCTTCTG 7709

QY 301 CGTGGGAGTCACTATGACACTATGACGACCACTGAACAGATGCTTGGTACACAA 360
    |||||||
DB 7710 CGTGGGAGTCACTATGACACTATGACGACCACTGAACAGATGCTTGGTACACAA 7769

QY 361 GTAAGCTCTGCACTTGTGAGCAGATGCTGCCGTCGGGTCATGTTTCACTTGTG 420
    |||||||
DB 7770 GTAAGCTCTGCACTTGTGAGCAGATGCTGCCGTCGGGTCATGTTTCACTTGTG 7829

QY 421 GATATTAAACAGGAATTTTATCTTGTGCAATGTTGTGATTTGTGAGGATAGTA 480
    |||||||
DB 7830 GATATTAAACAGGAATTTTATCTTGTGCAATGTTGTGATTTGTGAGGATAGTA 7889

QY 481 A 481
    |
DB 7890 A 7890

RESULT 5
LOCUS      D90807              20231 bp    DNA          linear    BCT 29-MAY-1997
DEFINITION E.coli genomic DNA, Kohara clone #316(36.7-37.1 min.).
ACCESSION  D90807 AB001340
VERSION    D90807.1 GI:1742692
KEYWORDS   Complete and shotgun sequencing; add; gsc; nth; pdxH; lys; ydha;
            yhcp.
SOURCE     Escherichia coli (strain:K12) DNA, clone_11b:Kohara lambda miniset
            library clone:Kohara clone #316.
ORGANISM   Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE  1 (sites)
AUTHORS    Aliba,H., Baba,T., Fujita,K., Hayashi,K., Inada,T., Isono,K.,

/codon_start=1
/transl_table=11

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS

COMMENT
The Japan E.coli genome DNA sequencing project
Project:
The Japan E.coli genome DNA sequencing project
Group:
The Japan E.coli genome DNA sequencing group
Members: (1995.4 - 1996.3)
Aliba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A.,
Horiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S.,
Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S.,
Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K.,
Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K.,
Nakamura,Y., Nashimoto,H., Nishio,Y., Oshima,T.,
Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K.,
Wada,C., Yamamoto,Y. and Yano,M.
Headed by:
Name: Takashi Horiuchi
Address: National Institute of Basic Biology, Okazaki, 444, Japan
E-mail: kishori@nibb.ac.jp
Information operator:
Name: Hirotsada Mori
Address: NARA Institute of Science and Technology,
Ikoma, 630-01, Japan
E-mail: hmo@legic.aist-nara.ac.jp
URL:
The Japan E. coli genome database
http://bsw3.aist-nara.ac.jp.
Location/Qualifiers
1.20231
/organism="Escherichia coli"
/strain="K12"
/db_xref="taxon:562"
/map="36.7 min"
/clone_11b="Kohara lambda miniset library"
/note="Nucleotide position 1705226-1725456 from the
Initiation site of Theta (0 min.). This clone is from
Kohara lambda miniset library"
1.840
/contig="add"
<1.840
/contig="add"
/note="ORF_ID:0314#9
similar to [SwissProt Accession Number P22333]"
/codon_start=1
/transl_table=11

```

/product="Adenosine deaminase (EC 3.5.4.4) (Adenosine
 amlnohydroase)".
 /protein_id="BAAL5388.1"
 /db_xref="GI:1742693"
 /translation="DLVSFLTKLDMGVKVLASIDACRRVAFENIEDAARHGILHYELR
 FSPGPMMAHDLPVAGVAVAVIDREGCTPFVQAGLIGISRTGEAAQCOIEAF
 LAHDDOITLADDELGPFGSLFSEHNRPAQGMHITVHAGEAGPSIMQAIHEL
 GAERIGHGVKRIEDRALMDPLAEQOIGIESCLTINQTSVLAELHAPLTKTFLEHIR
 ASINPDGQGVGDVTHIEITYAAFAAGLSREQIRQADINLEMAFLSAEKRALRKVY
 AAK"
 CDS
 complement(874..1953)
 /note="ORF_ID:0314#10
 similar to [SwissProt Accession Number P37168]"
 /codon_start=1
 /transl_table=1
 /product="viralence factor m7.m."
 /protein_id="BAAL5389.1"
 /db_xref="GI:1742694"
 /translation="MSNSHHTSMRGNDMSDNTRGVLIGYGYASTKEHAPLIACPGOE
 LHVYSSSEYKADMDPVTYVVSSEPKLFLNDPNLDIVTPMDTHPLAKALLEGK
 HAVVKKPTVTLISQARELDALAKLIGVTLSEHNRSDSPFLTKGLAGVLEGVAY
 PESHDREPPVDRPDRMGQSGSIVGLDALPLDDAITLFGIPVSMVDLQNLPG
 AOSTDYFAIISYQPRRIATLNGTLLAAESRYIVHRSRSGYKGLDPEEELNGE
 RLPQDDMDGYLTVNVESERYEETLIPGNTPAYTAAIRDALNGDENPVYAS
 QAIQVMEIIEGISAKHRAVTLCA"
 3954..4532
 /note="ORF_ID:0316#4
 similar to [SwissProt Accession Number P00211]"
 /codon_start=1
 /transl_table=1
 /product="Fettedoxin II."
 /protein_id="BAAL5390.1"
 /db_xref="GI:1742695"
 /translation="MNAIIVAAVAVSLIGLAFGAILGYASRRFAVEDDPVEKTIDEL
 POSQGGGCGPCGRPYEAISNGEKIRCAQGEAAVWLKAEILANEPDLGEOE
 ITPAMVAVIDNNQICGTICQACPDVAIVGATRAHHTVMSDLCGTCNLCVDPCTH
 CISLDQVAVETPDSMKMDLNTIPRIIPEHHA"
 4525..6747
 /note="ORF_ID:0316#5
 similar to [SwissProt Accession Number P14922]"
 /codon_start=1
 /transl_table=1
 /product="Glucose repression mediator protein."
 /protein_id="BAAL5391.1"
 /db_xref="GI:1742696"
 /translation="MLKLSAFRRKKIKVDENGCGTHPPBMTQSGCTPLRQVPLAQRVY
 IPIKSHIGELNLCVSGVDKVLRGQPLTRGKGKMLPVHAPSTGVTAIAPRSTAHPSA
 LAGELVITIDADEGCMIPRDGADRTSRREBELIERIHQFVAGLGGAGPTGKLOGG
 GGIKTIETLINAACEEPTITADDRIMQCAQVVEGILIAHLIIQPREIITIGIEDNP
 QAIISMRAKLVDSNDISLRVITPKYPSGAGAVIITLKGQPHGGSSDIQVADQW
 GRAYAKRAVADSGPIETERVYITGEALARGGNWAKRGVPHRLINDAGCGPMDOM
 VIMGELKGFETPLMDVPRVYKTNCLLAPRANLEGEDEBESCIRCSNADACADILL
 POOLWFSSGQOHDAATTHNIDATIEGCAQWVCNSITPLVOTRKEAELIARIKOE
 KRAAEKAFGEORARLEREKARLERIKSAVOPPAADKQAIALAALARKYEQAOXT
 OPYVIAGSRPNSAIIARAEARKQAQAELOQTMDAATVADPRTVIEAIAEA
 KARLEQOQANAEPEQVDPKRAVEAIIARAKARLEQOQANAEPEQVDPKRAVEA
 AIAARAKARLEQOQANAEPEQVDPKRAVEAIIARAKARREQDPAANAEPEQVDP
 KRAAEVAIARAKARLEQOQANAEPEQVDPKRAVAALAAARQAKKAQQKYNVED
 "
 7810..8430
 /note="ORF_ID:0316#10
 similar to [SwissProt Accession Number P44291]"
 /codon_start=1
 /transl_table=1
 /protein_id="BAAL5392.1"
 /db_xref="GI:1742697"
 /translation="MLKTRKHGIGITLALFAGSTGLTAIINQMTKTTIAEQASLQQRK
 LEPQVPAEYNNALAOQCYLITAPELQEGEHRVYIAQODRKYAIVLEAIPGYGSG
 AIDLLGAPNGVIGCTRVTEHNETPGIGDKTELRISMTIHPGKKISGADNDHAMVAV
 KRGGSDQDGTGTTIPRAVVAVAKRAGLRYOTPLAQLSQAPACE"
 8434..9129
 /note="ORF_ID:0316#11
 similar to [SwissProt Accession Number P43959]"

/codon_start=1
 /transl_table=11
 /protein_id="BAI15393.1"
 /db_xref="GI:1742698"
 /translation="MSEIKDYIYOGIMKNNLSALVOLGLCPILAVTSNATNALGLGLA
 TTVLALITLNTLITSTLRHWTPAEIRIPYIYMIATSVASVSAOMILINAYAGLDSIGITP
 PLIYINCIYVGRAEFAAKKGPALSLDIFSIGMAQCAVMPOLIRITIGNGLTFGSG
 ADALGSMKAVLYVEIFHTDPSFLMLRPPGAFIGLIMAGKYLIDERMKRRAEAA
 AERALPNSGETGNV"
 9129..9764
 /gene="nth"
 /note="orf_ID:0316#12
 similar to [PIR Accession Number A32412]"
 /codon_start=1
 /transl_table=11
 /product="Deoxyribonuclease (pyrimidine dimer) (EC
 3.1.25.1)"
 /protein_id="BAI15394.1"
 /db_xref="GI:1742699"
 /translation="IMKARKLEILITLRENNPHPTLELNFSSPELLIAVLISQATID
 VSNKATLLEIPVANTPAAMLELGEVGVYITTEIDLYSKSENIITKICRIILDHNGH
 EYDPERALALELPGVGRKTANVLTNATFAGPTITAVDTHTFRCNCTOFAPOKNEOVE
 EKLKLVPAEFKRVCHHMLILHGRYCIARKPRGSCITIEDCEYKEKVDI"
 10375..11877
 /note="orf_ID:0316#13
 similar to [Swissprot Accession Number P36574]"
 /codon_start=1
 /transl_table=11
 /product="Pi-cripeptide transporter."
 /protein_id="BAI15395.1"
 /db_xref="GI:1742700"
 /translation="MSTANOKPTFESVSLANFKOPAFYLIFSEITELMERGVYVLOGGIN
 AVYIKOLGMEADSTILFSSFSALYIGLVAIGWGLDVLGTGRIMHGLAIVLAG
 ALVNASMGADAGIYVGMGAALAVNGELFKAPSSLSSTCEKNDPRIDGAEFTWYISVVA
 IGFSPSMDIATPMLAAKYGWSYAFALISVGLITIVNFAFCORWKOYSGKDPDEFIIN
 RULLITLIIIVALLAIATMILHNOEVRMALGVAVGIIVIEGKEAFMKAGARRKMK
 VAFITMLIEIIVFVYSOMPTSLNFEFAIRNVEHSILGLAVEPOALPWEIITIGS
 ILAIIYKMKDITLPMPKFAFGMYMGSGAFLIILPLGAKFRASDAQIVSYMLVAYGLGL
 SIGETIMSGELAMVAQDLVQRLMGFTMGSEFLITMGANLIGIYVAGMAVAVDNTDLS

OY 421 GATATTAAACGAGATTATATCTGTCGATGTTGTCGATTGTCAGGATAGTA 480
|||||
Db 2339 GATATTAAACGAGATTATATCTGTCGATGTTGTCGATTGTCAGGATAGTA 2398
OY 481 A 481
|
Db 2399 A 2399

RESULT 6
AE005386
LOCUS
DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 5 of
290
ACCESSION AE005386 AE005174
VERSION AE005386.1 GI:12515609
KEYWORDS
SOURCE
ORGANISM Escherichia coli O157:H7 EDL933.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
AUTHORS Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Postal,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
Genome sequence of enterohemorrhagic Escherichia coli O157:H7
JOURNAL Nature 409 (6819), 529-533 (2001)
MEDLINE 21074935
PUBMED 11206551
TITLE
2 (bases 1 to 13259)
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Postal,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
Direct Submission
Submitted (22-OCT-2000) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
JOURNAL
TITLE
Location/Qualifiers
FEATURES
SOURCE 1..13259
/organism="Escherichia coli O157:H7 EDL933"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:155864"
/note="enterohemorrhagic"
118..1710
gene /gene="malX"
/note="22626"
118..1710
CDS /function="transport: Transport of small molecules:
Carbohydrates, organic acids, alcohols"
/note="Residues 1 to 530 of 530 are 99.43 pct identical to
residues 1 to 530 of 530 from Escherichia coli K-12 Strain
MG1655: B1621"
/codon_start=1
/transl_table=11
/product="PTS system, maltose and glucose-specific II ABC"
/protein_id="AG5610.1"
/db_xref="GI:12515610"
/translation="MTAKTAPKVTLMFEFOOLGKTFMLPVALLSFGIMLGSSLS
HDVITLPLVGNPVLOALFTWMSKIGSFAFSLPVMFCIALPLGLARENKGYAPAFGE
GVAYAMNLAVNFWLTKNGKILPTDDAIVKANNIOIILICISIDPDIIGAVIAGIIVWM
LHERHNIRLDPALAFEGGTRVPVPISSIVMGLVCLVPLVPIRAMKISGIGHINS
AGDGPMLFGTEGRLLDPLGHLIIIVALLRFDADAGTQECQIVSGALITFOALSC
PTTHGSSATFERSQGMKPAFLGLPGALALAMYCARPENKQIKGLISGLIACVY
GSTEPELEFLFVAPVLYVLIHALITGGLFTVMVSVIGTIGNTDNIIDFVYFGLHG
LSTKWMYPPVVAWIWFVYVYVIFRFAITRFNLKTPGRDSEVASSIEKAVAGAPKSGSY

gene
1720..2892
/gene="malY"
/note="22627"
1720..2892
CDS /gene="malY"
/function="enzyme: Degradation of small molecules: Carbon
compounds"
/note="Residues 1 to 390 of 390 are 99.23 pct identical to
residues 1 to 390 of 390 from Escherichia coli K-12 Strain
MG1655: B1622"
/codon_start=1
/transl_table=11
/product="enzyme that may degrade or block biosynthesis of
endogenous mal inducer, probably aminotransferase"
/protein_id="AG5611.1"
/db_xref="GI:12515611"
/translation="MEDESKYVDHRGTCQOWDYADBRGTADLLPTITSDMFATP
CITELNORLMHGVGRYRKNDERLAIAMHFTSTOHTAIDPOSVVGPVITAYSE
LIRQWSETEGVVHTPAYDAFYKAIENQRTVMVALEKQDGMFCDMGKLEAVLAK
PECKIMLCSQNPQKVTWCDELEIMADLERHGVRIYSDIHMDMYGEQPHIPWS
NVARQDMALITSGSKSFNIPALTGAVGIENSSDAYLSALKGHGLSPSLALTA
HIAVQOGAPMIDALRYVTKDNLVYIADKMNAAPPELMMQIPOSTYIAMLDLRPINID
DNALOKATIEQEKAIMGYTYGEGRGFVRLNACPRSKLEKYGAGLINAIRAVR"
2996..3997
gene /gene="add"
/note="22628"
2996..3997
CDS /function="add"
Salvage of nucleosides and nucleotides
/note="Residues 1 to 333 of 333 are 99.39 pct identical to
residues 1 to 333 of 333 from Escherichia coli K-12 Strain
MG1655: B1623"
/codon_start=1
/transl_table=11
/product="adenosine deaminase"
/protein_id="AG5612.1"
/db_xref="GI:12515612"
/translation="MIDTTLPLTDIHRHLDGNIRPOTILELGRQYNISLPAOSLETLI
PHYGVIAEPDLVSFLTKLDMGVYIASLACRVAFENIEDAANGLYHVLRFSPG
YMAAHOLPVGVGVAVIDGVREGRGTGYOAKLIGISRTGEAAQOALEAFAR
QDITADLDIAGDELGRPGSLFUSHENRARDAGHRTTVHNGEAAGPSINQOALELCAER
IGHGVKATIEDRALMDLAEQOIGIESCLTSNIQSTVVALEAHLAPLKTLEHETRASIN
ITDDPGVQGVDDIIEHYTVAPAGALSREQIQAQINGLEMAFLNAEKKALREKVAK"
complement(4033..5112)
/gene="22629"
complement(4033..5112)
CDS /gene="22629"
/function="orf: Unknown function"
/note="Residues 1 to 359 of 359 are 99.72 pct identical to
residues 1 to 359 of 359 from Escherichia coli K-12 Strain
MG1655: B1624"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AG5613.1"
/db_xref="GI:12515613"
/translation="MSNSHHTSRGNDMSDNIRVGLIGYAKTFEHLPLINGTGLE
LAVISSDETRKADMPVTVYVSEPKHLFNDPNIDLIYIPPNQDHPFLAKALRAGK
HYVNDKPTVTISQARELDALAKSGRVLVSFHNKRMSPDLITLGLALAEVLYGAVY
FESHFDRFPQVDRKWRDGGSGSIWYDLAHLDDATITTEGLPSVMTVDLAQLRGE
AOSTDYFHAIISYDQRRVYILHGTMLAAEASAYIVHSGSGYVKGLDPOERLRNPE
RLPOEDMGYDMRDGVLTVREGEEREVEETLLTPGNYPAYVYAIRDALNGDGENPVAS
QAIVQWELIEGIESAKRRATFCLA"
5715..5930
gene /gene="22631"
5715..5930
CDS /function="orf: Unknown function"
/note="Residues 1 to 71 of 71 are 100.00 pct identical to
residues 1 to 71 from Escherichia coli K-12 Strain


```

MG1655: B1625"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG5614.1"
/db_xref="GI:12515614"
/translation="MTVDYLKFKRKISLESLEKLYDHINLTLDQELLNMYRAD
HRRRLVSGRRFLDGLGVPKSVMHVQ"
5992..6456
/gene="2632"
/gene="2632"
5992..6456
/gene="2632"
/function="orf; Unknown function"
/note="Residues 1 to 154 of 154 are 98.70 pct identical to
residues 1 to 154 of 154 from Escherichia coli K-12 strain
MG1655: B1626"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG5615.1"
/db_xref="GI:12515615"
/translation="MRRYSGEPMTTTPQIRGMLGPIAMLIYALSTLLALYTA
ALSSPQFTLGAQALTTQILMGVSFTIALAMVYTLMLTIAFKRRRCYFKHTIML
LISVLAVKAFAPSPVEDGLAVRQLFTLLATLIVPFKRSSHVKATFVNP"
6533..7114
/gene="2633"
/gene="2633"
6533..7114
/gene="2633"
/function="orf; Unknown function"
/note="Residues 1 to 193 of 193 are 99.48 pct identical to
residues 1 to 193 of 193 from Escherichia coli K-12 strain
MG1655: B1627"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG5616.1"
/db_xref="GI:12515616"
/translation="MTDYLLFVGTVLVNNFVLEGLCPFMGVSKKLETAMGNGLA
TTPVWTASTICAMLDITWILPLNLTIRTAFLIVAVVQFTSMVVRKSPVLYRL
IGIRPLITTCNVIGVALLNINIGHNLQSLALGFSAAVGSLSLVMLFAAIRRLAV
ADVPAPFRGNALITLITAGLSLAEKFGVGLKL"
7114..7692
/gene="2634"
/gene="2634"
7114..7692
/gene="2634"
/function="orf; Unknown function"
/note="Residues 1 to 192 of 192 are 98.95 pct identical to
residues 1 to 192 of 192 from Escherichia coli K-12 strain
MG1655: B1628"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG5617.1"
/db_xref="GI:12515617"
/translation="MAIMIAVAAYXLLGLAFLGAILGYASRRFAVEDDPVYEKIDEL
FOSQCGQGPGRPYAIAISCKENKRCACGAGVAMKLAELLNVEPQIDGAE
LTPARMVAVIDENNICGCTICACPDVAICPATAMHTVMSDLCTGCNLCVDPCTH
CISLOPVAETPSMKMDLNTIPVRIIPEHNA"
7685..9907
/gene="2635"
/gene="2635"
7685..9907
/gene="2635"
/function="putative membrane; Not classified"
/note="Residues 1 to 740 of 740 are 97.83 pct identical to
residues 1 to 740 of 740 from Escherichia coli K-12 strain
MG1655: B1629"

```

```

Query Match          94.0%: Score 452.2: DB 1: Length 13259:
Best local Similarity 96.3%: Pred. No. 6,2e-108:
Matches 463: Conservative 0: Mismatches 18: Indels 0: Gaps 0:
QY 1 TGCCTCTCATTTGAAGTATGATGCTATTTGACACTATCCTTTACCCAGCTCAAGCTTT 60
|||||

```

```

Db 5079 TGCCTCTCATTTGAAGTATGATGCTATTTGACACTATCCTTTACCCAGCTCAAGCTTT 5138
QY 61 AATAACCTGCCAGCATTAAGGATGTTGTTAACTTAAGTCAAAAAATAGCAATTTTC 120
|||||
Db 5139 AATAACCTGCCAGCATTAAGGATGTTGTTAACTTAAGTCAAAAAATAGCAATTTTC 5198
QY 121 CAACGACAAAAGCTAAATATGCAAAAACCTCAGTAAATCTTGGTGGACATTAATT 180
| | | | |
Db 5199 CCAACGACAAAAGCTAAATATGCAAAAACCTCAGTAAATCTTGGTGGACATTAATT 5258
QY 181 GCTAAGTAACTTTTACCCCTGAAGTTAATGATCAATCAAGATGAGGCTGTAAT 240
| | | | |
Db 5259 GCTAAGTAACTTTTACCCCTGAAGTTAATGATCAATCAAGATGAGGCTGTAAT 5318
QY 241 GAATCGCTTTATTAATTAACAGGTGGATGCTTGTGCTTTCAGTCAATCTTCTTGG 300
| | | | |
Db 5319 GAATCGCTTTATTAATTAACAGGTGGATGCTTGTGCTTTCAGTCAATCTTCTTGG 5378
QY 301 CGTGCGAGTCACATTTGACAACTATCAGCCACCTGACAGAGTCTTGGTACACAA 360
| | | | |
Db 5379 CGTGCGAGTCACATTTGACAACTATCAGCCACCTGACAGAGTCTTGGTACACAA 5438
QY 361 GTAACTCTGCATTTGGAGCGAGCATGCTCCGCTCCGGTGCATGTTTCACTTGCG 420
| | | | |
Db 5439 GTAACTCTGCATTTGGAGCGAGCATGCTCCGCTCCGGTGCATGTTTCACTTGCG 5498
QY 421 GATATTAACACGAAATTTATTCCTTGTGATGTTGTTGGTATGTCAGGATAGTA 480
| | | | |
Db 5499 GATATTAACACGAAATTTATTCCTTGTGATGTTGTTGGTATGTCAGGATAGTA 5558
QY 481 A 481
Db 5559 A 5559

RESULT 7
AP002558
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Escherichia coli O157:H7
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
AUTHORS
Makino,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S.,
Kurokawa,K., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T.,
Yamamoto,K., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T.,
Sasakawa,C. and Shinagawa,H.
Complete nucleotide sequence of the prophage VT2-Sakai carrying the
verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
derived from the Sakai outbreak
Genes Genet. Syst. 74 (5), 227-239 (1999)
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
Comparative analysis of the whole set of rRNA operons between an
enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
Escherichia coli K-12 strain MG1655
Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
Complete nucleotide sequence of the prophage VT1-Sakai carrying the
Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli

```


/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB35727.1"
/db_xref="GI:13361771"
/translation="MRTIIVLGIISLTSFVFPADKPDVAKSANDEVSTLFFGHDR
VSNKGLWYDIDHIDEGRDVPTLGRKLRKADGDMIVPPAAMPDFSLIVLRNPSSGTP
LPLYREIKPRLEPH"
/complement(7657..7986)
/gene="ECs2305"
/complement(7657..7986)
/gene="ECs2305"
/note="similar to B1599_ECOLI g1178783 percent identity
100 in 109 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence-not_experimental

Query Match 93.5%; Score 449.8; DB 1: Length 257071;
Best Local Similarity 97.3%; Pred. No. 2.3e-107;
Matches 468; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 TGGCTCTCATTTGAATGATGATGCTATTTGACACTATCTTTACCCAGCCTCAACAGTTT 60
DB 38981 TGGCTCTCATTTGAATGATGATGCTATTTGACACTATCTTTACCCAGCCTCAACAGTTT 39040
QY 61 AATAACCTGGCCAACTAAGGATGTTGTTAACTTAAGTCAAAAATAATGCGAATTTTC 120
DB 39041 AATAACCTGGCCAACTAAGGATGTTGTTAACTTAAGTCAAAAATAATGCGAATTTTC 39100
QY 121 CAACGAACAAAGCTAAATATGCAAAACCTGAGTAAATCTGTGTGAGCTATTATT 180
DB 39101 CAACG-ACAAAGCTAAATATGCAAAACCTGAGTAAATCTGTGTGAGCTATTATT 39159
QY 181 GCTAAGTAACTTTACCCCTGGAAGTTAATGATCAATCAAGAGATGTGGCTGTAAT 240
DB 39160 GCTAAGTAACTTTACCCCTGGAAGTTAATGATCAATCAAGAGATGTGGCTGTAAT 39219
QY 241 GAATCGTCTTATTTGAATTAACAGTTGATGCTCTTGTGCTTCAATCTTCTTGG 300
DB 39220 GAATCGTCTTATTTGAATTAACAGTTGATGCTCTTGTGCTTCAATCTTCTTGG 39279
QY 301 CGTGGCAGTCACTTGAACATATGACCACTGACAGACAGATGCTTCCGTGTAACACAA 360
DB 39280 CGTGGCAGTCACTTGAACATATGACCACTGACAGACAGATGCTTCCGTGTAACACAA 39339
QY 361 GTAAGCTCTGCACTTGTGAGCAGATGCTGCCGCTCCGGGTCAGTGTTCACCTGTG 420
DB 39340 GTAAGCCTGCGCTTGTGAGCAGATGCTGCCGCTCCGGGTCAGTGTTCACCTGTG 39399
QY 421 GATATTAACACGAAATTTATCTGTGTCAGATGTTGGATGTCAGAGGATGTA 480
DB 39400 GATATTAACACGAAATTTATCTGTGTCAGATGTTGGATGTCAGAGGATGTA 39459
QY 481 A 481
DB 39460 A 39460

RESULT 8 358 bp DNA linear BCT 09-AUG-2000
AF219227
LOCUS Escherichia coli beta-lactam resistance protein (blr) gene,
DEFINITION complete cds.
ACCESSION AF219227
VERSION AF219227.1 GI:8953940
KEYWORDS Escherichia coli.
SOURCE Escherichia coli.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
REFERENCE 1 (bases 1 to 358)

AUTHORS Wong, R.S., McMurtry, L.M. and Levy, S.B.
TITLE 'Enterogenic' blr gene in Escherichia coli encodes a 41-residue
membrane protein affecting intrinsic susceptibility to certain
inhibitors of peptidoglycan synthesis.
JOURNAL Mol. Microbiol. 37 (2), 364-370 (2000)
MEDLINE 10931331
PUBMED 20392465
REFERENCE 2 (bases 1 to 358)
AUTHORS Wong, R.S.Y., McMurtry, L.M. and Levy, S.B.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Molecular Biology & Microbiology, Tufts
University School of Medicine, 136 Harrison Ave., Boston, MA 02111,
USA

FEATURES
source Location/Qualifiers
1..358
/organism="Escherichia coli"
/sub_species="K12"
/db_xref="taxon:562"
/map="min 36.6"
93..>257
/gene="blr"
/note="similar to corresponding region of Escherichia coli
sequences deposited in AB00258 and D90870"
93..>257
/gene="blr"
/product="beta-lactam resistance protein"
132..257
/note="new annotation of protein Blr"
/codon_start=1
/transl_table=11
/product="beta-lactam resistance protein"
/protein_id="AAF82191.1"
/db_xref="GI:8953941"
/translation="MNRILELGMIVLVSVILLGVASHIDNYQPEQSASVQHK"

BASE COUNT 100 a 71 c 78 g 109 t

Query Match 72.1%; Score 347; DB 1: Length 358;
Best Local Similarity 99.7%; Pred. No. 2.6e-80;
Matches 358; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 107 AATAGCGAATTTTCCACGACAAAGCTTAATATGCGCAAAAACCTCATTAATAATCTTG 166
DB 1 AATAGCGAATTTTCCACG-ACAAAGCTTAATATGCGCAAAAACCTCATTAATAATCTTG 59
QY 167 CTGGAGCTATTTATGTAAGTATACATTTACCCCTGAAGTTAATGATCAATCAAGAGAG 226
DB 60 CTGGAGCTATTTATGCTAAGTATACATTTACCCCTGAAGTTAATGATCAATCAAGAGAG 119
QY 227 ATGTGGCTGTATGAATGCTTTATTTGAATTAACAGTTGATGCTTCTGTGCTTTCA 286
DB 120 ATGTGGCTGTATGAATGCTTTATTTGAATTAACAGTTGATGCTTCTGTGCTTTCA 179
QY 287 GTCAATCTTGTGGCGTGGAGTCAATGACAACTATCAGCCACTGACAGAGTGTG 346
DB 180 GTCAATCTTGTGGCGTGGAGTCAATGACAACTATCAGCCACTGACAGAGTGTG 239
QY 347 TCGGTACACACAAAGTAAAGCTTGCATCTGTGAGGACACATGCTGCCGTCCGGGTGCAT 406
DB 240 TCGGTACACACAAAGTAAAGCTTGCATCTGTGAGGACACATGCTGCCGTCCGGGTGCAT 299
QY 407 GTTTTCACTTGTGGATATTAACCGAAGATTTATATCTTGTTCAGATGTTGTGCTGA 465
DB 300 GTTTTCACTTGTGGATATTAACCGAAGATTTATATCTTGTTCAGATGTTGTGCTGA 358

RESULT 9 19211 bp DNA linear BCT 29-MAY-1997
D90805
LOCUS D90805
DEFINITION E.coli genomic DNA, Kohara clone #314(36.3-36.7 min.).
ACCESSION D90805 AB001340
VERSION D90805.1 GI:1742665

KEYWORDS Complete and shotgun sequencing; add: funa; func; hdb; hsdh; mali; maly; maly; mana; pmi; uida; gusa; gura; uida; gusb; uidp; yihf.

SOURCE Escherichia coli (strain:K12) DNA, clone_id:Kohara lambda miniset library clone:Kohara clone #314.

ORGANISM Escherichia coli

REFERENCE Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (sites)

AUTHORS Aiba,H., Baba,T., Fujita,K., Hayashi,K., Inada,T., Isono,K., Itoh,T., Kasai,H., Kashimoto,K., Kimura,S., Kitagawa,M., Kitagawa,M., Makino,K., Miki,T., Mizobuchi,K., Mori,H., Mori,T., Motomura,K., Nakade,S., Nakamura,Y., Nishio,Y., Oshima,Y., Oshima,T., Saito,N., Sempel,G., Seki,Y., Sivasubraman,S., Tagami,H., Takeda,J., Takemoto,K., Takeuchi,Y., Wada,C., Yamamoto,Y. and Horiuchi,T.

TITLE A 570-Kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map

JOURNAL DNA Res. 3 (6), 363-377 (1996)

MEDLINE 97251357

REFERENCE 2 (sites)

AUTHORS Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitagawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishio,Y., Oshima,T., Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.

TITLE The systematic sequencing of the Escherichia coli genome in Japan

JOURNAL Unpublished

AUTHORS 3 (bases 1 to 19211)

TITLE Direct Submission

JOURNAL Submitted (14-DEC-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5

REFERENCE Takayama, Ikoma, Nara 630-01, Japan

AUTHORS (E-mail:hmori@gtc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)

TITLE Collaboration Information:

JOURNAL Project:

COMMENT The Japan E.coli genome DNA sequencing project

REFERENCE Group:

AUTHORS The Japan E.coli genome DNA sequencing group

TITLE Members: (1995.4 - 1996.3)

JOURNAL Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitagawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishio,Y., Oshima,T., Saito,N., Sempel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.

HEADED BY:

NAME: Takashi Horiuchi

ADDRESS: National Institute of Basic Biology, Okazaki, 444, Japan

E-MAIL: kishori@nibb.ac.jp

INFORMATION OPERATOR:

NAME: Hirotsada Mori

ADDRESS: NARA Institute of Science and Technology, Ikoma, 630-01, Japan

E-MAIL: hmori@gtc.aist-nara.ac.jp

URL:

THE JAPAN E. COLI GENOME DATABASE

HTTP: bsw3.aist-nara.ac.jp.

LOCATION/QUALIFIERS

1. 19211

FEATURES

SOURCE /organism="Escherichia coli"

/db_xref="taxon:562"

/map="36.3 min"

/clone="Kohara clone #314"

/clone_lib="Kohara lambda miniset library"

/note="Nucleotide position 1688148-1707358 from the initiation site of ThnA (0 min.).-This clone is from

gene Kohara lambda miniset library"

complement(1. .1272)

/gene="fuma"

complement(<1. .1272)

/gene="fuma"

/note="ORF ID:0312#11

similar to [PIR Accession Number S07138]"

/codon_start=1

/transl_table=11

/product="Fumarate hydratase (EC 4.2.1.2) FumC"

/protein_id="BAAL5363.1"

/db_xref="GI:1742666"

/translation="MNTVSEKDSKALIDVPADKIMGATQTSLEFRISTEKPTSL IHALATKRAAKVNEEDUGLSEKASATROADEVLACQHDDEFLATWQSGTOS NNNMNEVLNARSSELLGVGRMERKVPDDVNVKSSQSDVPEPTAMHVALLRKOL IQPLRTDTLNEKSRADFIVKIGTHLDADATPLTGOEISGWAMLEHNLKHEYS LPHVAELAGTAVGTLNTHPEYARVADDELAVITCAPFVAPNFEALATCDALVQ AHGALKGLAASLMKTIANDVRLWASGRCGISTISPENRGSSIMPGKVNPCCEALT MICCOVMGNDVAIIMGASGNEFLNFRPMVTHNFTQSVRLADGDSFNKCAVGE PRERINQNLNLSMLVTLNTHIGDKAAET"

complement(1415. .3061)

/gene="fuma"

complement(1415. .3061)

/gene="fuma"

/note="ORF ID:0312#12

similar to [PIR Accession Number A03531]"

/codon_start=1

/transl_table=11

/product="Fumarate hydratase (EC 4.2.1.2) Fuma,

iron-dependent"

/protein_id="BAAL5364.1"

/db_xref="GI:1742667"

/translation="MSNKPFFHQAPEPLKKDTEYLLTSEHVSSEFEQELIKVAP EALLTLARFHDASFMRLPAHQOQVADILRPEASENKRYYALQFLRSDIANKCVL PICQDTGTAIVIGKQGVATGGDEAALARGYVNTIDNLTNSYNSAPLDWKKEVNT GTNLPADQILVAVDGEYKELCIKAGGSANKTYLQETRKALLETPEKLTNYLEKERT LGTACAPPHVIAFVIGGTSAGETNKLKVKASAKYVDLPENGEHCAARDVLEKEL LLEAONLIGLGAOPGGKYFPHDIRVILPRHGASCPYMGKVSQADNITAKIRGQIW IERLEHNPGRYIPEELRKAGEAVVNDLNRPKETLQLOSOVPSTRLSNGTIIIG RPIAHAKLERMDNEGSLPOYIKDHPYIYAGPAKTEGYASGSLGPTGTGRDSYVD LGAOGSMIMLAKGNSQOVTPACKKHGFGYLSIGPAVAVLAGSISKELECYEPEL GMAIKWIEVEDPEPFIIVDKNDFQQLQTLQCRCVK"

3260. .4435

/gene="mana, pmi"

3260. .4435

/note="ORF ID:0312#13

similar to [SwissProt Accession Number P00946]"

/codon_start=1

/transl_table=11

/product="Mannose-6-phosphate isomerase (EC 5.3.1.8) (Phosphomannose isomerase) (PMI) (Phosphohexomutase)."

/protein_id="BAAL5365.1"

/db_xref="GI:1742668"

/translation="MOKLINSQNVYAMGSKTALTEYGHENPSSQPMALMCAHNRKS SSRVQNAAGDIVSLRDVTSSTLGEVNAKPGELPLFLKYLCAQPSLSTOVHNRK HNSEIGFAENAGIPMDAERNYKDPNKRPELVALTPEFLANARRESEYSLIOP VAGHAPAIAPHILOQDPDAERLSEFASLIMMGQEKSRALAILKSLADSOOGFEWQIR LSEFYPEDSGLEFPLLNVLNVKLNPEAMLEFETGHAVILOGVALVPMNSDNRVLAG LIPKYIDIPELVANYEFAKPPANOLLTQPVKQAGELDPIVPDDFAFSIHLSDSKETT ISQGSAAITFCYBGDXTLMKGSQQLDKGESAFLIANSPYVYKGHRLARYNKL"

4536. .6044

/gene="yihf"

4536. .6044

/note="ORF ID:0312#14

similar to [SwissProt Accession Number P32128]"

/codon_start=1

/transl_table=11

/protein_id="BAAL5366.1"

/db_xref="GI:1742669"

/translation="MNRSLVAVGCVIYALGVVMTGAMYGKKIETHELEDMVQAANQL KITAPESNLEVSQNTNRGVFSQQLLVKPIAGKNEPWIKSGSVIFNESVDHGFP

```
LAQKRLNLPMSASIQTLVNNVSKPLFDMAGETPEELNSITGSCSSDISLK
PLNVEOKDEKVAFGSGEPOLNADRDGKAISLSGAOSGRIDAVNEYKQVLTNNLK
TDSSTLASFGERNOKLSLEKMTISVEKELALLEMEISGSDIADNDKTI NSOL
DYSLSLKVONODGSKLTLKVGQIDSEAHNHOSSQYNAQTOALADPELANNBELY
OEKTEAFESALPLMLKGDPAVITTAIPGSKMSOGSEALNLSLFDKPTTEAEOTLA
OEVDRSVSLDAKLITPYDMA TERMTQAKLEGQOEODAKLALQOVSASAMQMR
LTTTODNTTTSLOITANGQITLNGOKSLDEPVGMEFAPALNVAVPAIPQO"
/gene="uidb, gusb, uidp"
/complement(7574..8947)
/gene="uidb, gusb, uidp"
/complement(7574..8947)
/gene="uidb, gusb, uidp"
/notes="ORF_ID:0314#2
similar to [SwissProt Accession Number P30868]"
/codon_start=1
/transl_table=11
/product="Glucuronide carrier protein (Glucuronide
permease)."
/db_xref="GI:1742670"
/db_xref="GI:1742671"
/translation="NQQLSWRTIVGSLGDVNAFNAFMALFILSYTVDVAGVAA
AAGTLLVRFADADVPAGRVDSVNTRMKRPFLFETAPLMTISVLFVYVLTLD
WSHGSKVYAYLTVMGICYSLVNIPYGLATAPQDSARLGAARGIAASLTFV
CLAFLIGSIKNSPEEMVSVHFWITVLAAGVLYFICEKSTRENNVRIYAQSLN
ISLOTLRNRPLFMLCIGALCVLSTFPAVSASLFFVRLNDGLFTVLVYQMLVG
TVASAPVPGMVARIGKNTPLIGALIGTGSGLFFPWSVSLPALVALAISGGG
VMYVMALADPTVEXGTLGVRTGREGITLSLFFSTRKCGAGIGSIPAFILGSGYI
ANQVTPPEYIMKIRTSIALVPCGFMFLAFVITWIFPLDKKREIVEIDNNKKVQOQ
LISDITN"
/complement(8944..10755)
/gene="uida, gusa, gura"
/complement(8944..10755)
/gene="uida, gusa, gura"
/notes="ORF_ID:0314#3
similar to [SwissProt Accession Number P05804]"
/codon_start=1
/transl_table=11
/product="b-glucuronidase (EC 3.2.1.31) (GUS)
(b-d-glucuronoside glucuronosylhydrolase)."
/protein_id="BA15368.1"
/db_xref="GI:1742671"
/translation="MLRPVETPREIKRIDGLMARSIDRENCGIDORMWESALQESRA
IAPGSFNDQFADADIRNAGVNWTOREFIRGMAQGRIVLRFDAVTHRGVWVANO
EYMEHGGATPEADVTPTYIAGKSVIRITVCVNNELNMQITPEGMVITDENKRSQSY
FHDFFNYAGIHSRSMVLYTTPNTWVDITVTVHQAQDNASVDQWVANGADVSELAR
ADQOVVAAGCGTSGTLOVNPMLQPEGSLYELCYTAKSQCEDIDYPLRVGIRSVYA
KGEQFLINHKRPYFTGSEGRHEDADLRGKGDVNLVHDAKNDMGANGANSRSHRYA
EEMLDADENHGIVIDEPTAANGENSLIGTFAGNPKREKLEAEVANGFTQOAHQAI
KELIARDKNHPSVYMSIANEPRTPOGARRETFAPLAEXTKRILDEPRPTCYNNMCD
AHTDTISDLEFDVLCINRYIGWYVQSDLETAELVLEKELIAMQERKLHOPITITTEGYD
TLAHLMSYTDMSSEYQCAMLDMYHRVDFRVSAYVGEQVMWFADPVTISQILRVGNN
```

```
Query Match 42.2%; Score 203; Db 1; Length 19211;
Best Local Similarity 99.5%; Pred. No. 1.1e-42;
Matches 214; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 1 TGCCTCTCATTTGAAGTATGATGCTATTGACACTATCTTTACCCACGCTCAACAGTTT 60
|||||
Db 18998 TGCCTCTCATTTGAAGTATGATGCTATTGACACTATCTTTACCCACGCTCAACAGTTT 19057
|||||
QY 61 AATAACCTGCCAGCAATTAAGGATGCTGTTAACTTAAGCAAAAAAATAGCGAATTTTC 120
|||||
Db 19058 AATAACCTGCCAGCAATTAAGGATGCTGTTAACTTAAGCAAAAAAATAGCGAATTTTC 19117
|||||
QY 121 CAACGAACAAAGCTAATATATCGCAAAACCTAGTAAATATCTTGGCTGAGCATATTATT 180
|||||
Db 19118 CAACG-ACAAAAGCTAATATATCGCAAAACCTAGTAAATATCTTGGCTGAGCATATTATT 19176
|||||
QY 181 GCTAAGTAAACATTTACCCCTGGAAGTTAATGATC 215
|||||
Db 19177 GCTAAGTAAACATTTACCCCTGGAAGTTAATGATC 19211
|||||
RESULT 10
```

```
AE008763/c
LOCUS AE008763 23009 bp DNA linear BCF 31-JUL-2002
DEFINITION Salmoneella typhimurium LT2, section 67 of 220 of the complete
genome.
ACCESSION AE008763 AE006468
VERSION AE008763.1 GI:16419969
KEYWORDS
SOURCE Salmoneella typhimurium LT2.
ORGANISM Salmoneella typhimurium LT2.
REFERENCE 1 (bases 1 to 23009)
AUTHORS McClelland,M., Sanderson,K.E., Spieth,J., Clifton,S.W.,
Latreille,P., Courtney,L., Porwollik,S., Ali,J., Dante,M., Du,F.,
Hou,S., Layman,D., Leonard,S., Nguyen,C., Scott,K., Holmes,A.,
Grewal,N., Mulvaney,E., Ryan,E., Sun,H., Florea,L., Miller,W.,
Stonking,T., Nhan,M., Waterston,R. and Wilson,R.K.
Complete genome sequence of Salmoneella enterica serovar Typhimurium
LT2
JOURNAL Nature 413 (6858), 852-856 (2001)
MEDLINE 21534948
PUBMED 11677609
REFERENCE 2 (bases 1 to 23009)
AUTHORS The Salmoneella typhimurium Genome Sequencing Project.
JOURNAL Direct Submission
Submitted (29-MAR-2001) Genome Sequencing Center, Department of
Genetics, Washington University School of Medicine, 4444 Forest
Park Boulevard, St. Louis, MO 63108, USA
COMMENT Supported by NIH grant 5U 01 AI43283
```

Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs: GLIMMER: <http://www.tigr.org/softlab/glimmer/glimmer.html> and Genemark: <http://opal.biology.gatech.edu/Genemark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes: <http://www.genome.ad.jp/kegg/>, and Pedro Romero and Peter Karp at EcoCyc: [http://ecocyc.org/ecocyc/](http://ecocyc.org/ecocyc/PangeaSystems.com/ecocyc/)

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and RegionDB: http://kinich.cfin.unam.mx:8650/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m33 subclone.

```
FEATURES
source
location/Qualifiers
1..23009
/organism="Salmoneella typhimurium LT2"
/strain="LT2; SCSG 700720"
/db_xref="ATCC:700720"
/db_xref="taxon:99287"
/notes="LT2"
88..1374
/gene="tyrS"
/notes="STM1449"
88..93
/gene="tyrS"
/notes="putative RBS for tyrS; RegulonDB:STMS1H001673"
100..1374
/gene="tyrS"
/BC_number="6.1.1.1"
/notes="similar to E. coli tyrosine tRNA synthetase
(AAC74709.1); Blastp hit to AAC74709.1 (424 aa), 96%
identity in aa 1 - 424"
/codon_start=1
/transl_table=11
/product="tyrosine tRNA synthetase"
```

| | | | |
|---------------------------|---|--|---|
| RBS | 11AAIYVRMGDTLPMPKRFALGNVLCGATLLPLEAKFENDKAGIYSVNNLIASYGLO SIELEMTISGLIAMAVALPQRLMGFTMSWFLTTAGANIIIGGVANLMAVPSDYDTP LMSLEMYGRVEMQIGITAVIAVIMLTATKLNRMTODDDFAEKSKAAIV" | complement (4572) .4577) | |
| gene | /note="putative RBS for ydgR; RegulonDB:STMS1H001676" | complement (5170) .3818) | |
| CDS | /gene="nth" /note="STM1453" complement (5170) .5805) | /gene="nth" /EC.number="4.2.99.18" | /note="similar to E. coli endonuclease III; specific for apurinic and/or apyrimidinic sites (AAC74705.1); Blastp hit to AAC74705.1 (231 aa), 95% identity in aa 1 - 211; DNA glycosylase/aprilymidinoc (Ap) lyase" |
| | /codon_start=1 /transl_table=11 /product="endonuclease III" | /protein_id="AAL20375.1" | /db_xref="GI:16419974" |
| | /translation="MNKARKLETLRLRDNPNPPTLENTSPRELLIIVLSQAND VSVNRKAKIKIPIYANTPRAMLELGEVSKYIKTIGLFSNKAENVITGCIILDKING EKLEVDRALEALPGVGRKTIANVYLVNTAGWPTIAVDHIFRCVCRTOFAPGKNEVE EKLIVPNEFKYDCHMILLHGKRYTCIARRKPGSCLLIEDLEYKKEKVDI" | complement (5805) .6510) | |
| gene | /gene="ydgQ" /note="STM1454" | complement (5805) .6497) | |
| CDS | /gene="ydgQ" /note="similar to E. coli orf, hypothetical protein (AAC74704.1); Blastp hit to AAC74704.1 (231 aa), 91% | identity in aa 1 - 231" | |
| | /codon_start=1 /transl_table=11 /product="putative membrane NADH:quinone reductase protein" | /protein_id="AAL20376.1" | /db_xref="GI:16419975" |
| | /translation="MSEIKRIYVOGLKNNNSALVOLIGCPILAVSTATNALGLGIA TTLVLTINLTVALRKWTPAEIRIPIYVIMIASVAVMLINAAVEIGYOSIGIT PLIVNTCIVVGRAEAFPAKGPWLSALDPSIGSAVGAAMFVGLSLEIIGNOTLEDGQ ADSLTIGMAVULVRLVEIRFHTDSPFLAMLPGAFGLGLMLAVYLLDEKKMKRRARETA PSAIPAGETGKV" | complement (5813) .5818) | |
| RBS | /gene="nth" | /note="putative RBS for nth; RegulonDB:STMS1H001677" | |
| gene | complement (6500) .7131) | /gene="ydgP" | /note="STM1455" |
| CDS | complement (6500) .7120) | /gene="ydgP" | /note="similar to E. coli orf, hypothetical protein (AAC74703.1); Blastp hit to AAC74703.1 (206 aa), 84% |
| | identity in aa 1 - 206" | | |
| Query Match | 22.2%; Score 106.6; DB 1; Length 23009; | | |
| Best Local Similarity | 70.68; Pred. No. 2e-17; | | |
| Matches 156; Conservative | 0; Mismatches 64; Indels 1; Gaps 1; | | |
| OY 146 | AAAACTCAGTAAAAATCTGCTGGAGCTATTATTGCTAAGTAACATTATCCCGCTGAAG | 205 | |
| Db 12878 | AAAACGTTCCCTCCCTCTTGCCGACCGTTGCGCAGGAAAGTAACATTATACCCCTG-TG | 12820 | |
| OY 206 | TTAATGCATCAATCAAGAGAGATGGGGCTGTAATGAATGCTTTATTGAATTAACAGT | 265 | |
| Db 12819 | TTTATGCAATATCAAAACCAATGCGCAATATGAATGCTTTATTGAATTAACCTGGG | 12760 | |
| OY 286 | TGGAATCGTTCTTGCGTTTCAGTCATTTCTTGCGCGTGGGAGAGACATGACAACAT | 325 | |
| Db 12759 | TGGAATGTTCTGCTTATCTAGTATCTTTTATAGGTAATGCGCAACCATATATGACAATTAAC | 12700 | |
| OY 326 | CAGCCACCTGAAGAGAGTGTTCGGTACCAACACAAGTAAG | 366 | |

Db 12699 CAACCACGAGCCACGCGCTTCGTCGACAGAGTAAGC 12659

RESULT 11
AL627271 233050 bp DNA linear BCT 06-JUN-2002
LOCUS Salmonella enterica serovar Typh1 (Salmonella typh1) strain CT18,
DEFINITION complete chromosome; segment 7/20.
ACCESSION AL627271 AL513382
VERSION AL627271.1 GI:16502733

SOURCE
ORGANISM Salmonella enterica subsp. enterica serovar Typh1.
Salmonella enterica subsp. enterica serovar Typh1.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.

REFERENCE
AUTHORS 1 (bases 1 to 233050)
Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D.,
Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G.,
Sebatilla,M., Baker,S., Basham,D., Brooks,K., Chillingworth,T.,
Comerton,P., Cronin,A., Davis,P., Davies,R.M., Dowd,L., White,N.,
Farrair,J., Feltwell,T., Hamlin,N., Haque,A., Hien,T.T., Holtzcl,S.,
Jagels,K., Kirogh,A., Larsen,T.S., Leather,S., Moule,S., O'Gaora,P.,
Parry,C., Quail,M., Rutherford,K., Simmonds,M., Skelton,J.,
Stevens,K., Whitehead,S. and Barrrell,B.G.
Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typh1 CT18
Nature 413 (6858), 848-852 (2001)
JOURNAL MEDLINE
PUBMED 21534947
11677608

REFERENCE
AUTHORS 2 (bases 1 to 233050)
Parkhill,J.
Direct Submission
Submitted (25-OCT-2001) Submitted on behalf of the Salmonella
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, UK
E-mail: parkhill@sanger.ac.uk

COMMENT
Notes:
Details of S. typh1 sequencing at the Sanger Centre are available
on the World Wide Web.
(URL: http://www.sanger.ac.uk/projects/S_typh1/).
Location/Qualifiers
1. 233050
/organism="Salmonella enterica subsp. enterica serovar
Typh1"
/strain="CT18"
/db_xref="taxon:90370"
199. 1581
/gene="STY1645"
199. 1581
/gene="STY1645"
/note="Similar to Pseudomonas aeruginosa
arginine/ornithine antiporter acd SW:ARCD_PSPA6 (P18275)
(482 aa) fasta scores: E(): 0, 45.5% id in 479 aa
fasta hit to YDGE_ECOLI (445 aa), 31% identity in 450 aa
overlap
orthologue of E. coli ARCD_ECOLI; fasta hit to ARCD_ECOLI
(460 aa), 93% identity in 460 aa overlap"
/codon_start=1
/transl_table=11
/product="putative amino acid permease"
/protein_id="CAD01890.1"
/db_xref="GI:16502734"
/db_xref="SPRMBL:08XG1"
/translation="MKKIGLSLTLVLSVSMIGAGVSLPONMAVAPASLLTGM
ITGVGILLAFAMILTRIRSELDGSIIFYARGFEBLLIGFCANQWICAVIANYS
LVYFSAISFDTPALRLEGDNTQAIAGVALLMIHFLIRGVQPAASINLAT
LAKLPLGAFIVAIIMFKDFTLDFTEIGSIPIWEOVKNTMLITLWVIGVEAV
VVSARAKNRDVGATLLAVIALGIVLYLTSLGVALRPELAERNPSMAGLAKM
MGWGEITTAGIIVSCAYISLWTTMAAPVPLIAYSPFELFARONAGAPASL
WLNINVCQICLVIMLTGSDYNTLLTIASMLIVPFLVGAFLLKATRLPHRAVIG
ACTYGLMLVYASGPMHLLSVLYAPGLVFLVYARTHKHDNVLNQREVLLIGLLVA
AVPATMNLVG"

misc_feature 199. 1563
/gene="STY1645"
/note="Pfam match to entry PF00324 aa_permeases, Amino
acid permease, score -290.50, E-value 0.062"
complement(1643. 1978)
/gene="STY1646"
complement(1643. 1978)
/gene="STY1646"
/note="Orthologue of E. coli YDGC_ECOLI; fasta hit to
YDGC_ECOLI (111 aa), 81% identity in 111 aa overlap"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD01891.1"
/db_xref="GI:16502735"
/db_xref="SPRMBL:08Z6S2"
/translation="MGLYIKALGALVYVVLIGLSKTKNYIAGLIPFPFLLIAY
IVASERIDAKRTIVFSMWSIIFYLTATMTFSGVMLPYALGAGAVCKGLSAML
LIFCMIKWH"
2105. 2836
/gene="STY1647"
/note="rsta"
2105. 2836
/gene="STY1647"
/note="STY1647"
/note="Similar to Staphylococcus aureus response regulator
YycF yycF TR:09XCM7 (EMBL:AF136709) (233 aa) fasta scores:
E(): 8.8e-28, 37.6% id in 234 aa
fasta hit to TORR_ECOLI (230 aa), 31% identity in 213 aa
overlap
fasta hit to BAER_ECOLI (240 aa), 34% identity in 229 aa
overlap
fasta hit to YLCA_ECOLI (227 aa), 32% identity in 232 aa
overlap
fasta hit to YEDW_ECOLI (239 aa), 34% identity in 233 aa
overlap
fasta hit to OMPR_ECOLI (229 aa), 33% identity in 235 aa
overlap
fasta hit to CPXR_ECOLI (232 aa), 37% identity in 233 aa
overlap
orthologue of E. coli rsta (RSTA_ECOLI); fasta hit to
RSTA_ECOLI (242 aa), 85% identity in 239 aa overlap"
/codon_start=1
/transl_table=11
/product="putative two-component response regulator"
/protein_id="CAD01892.1"
/db_xref="GI:16502736"
/db_xref="SPRMBL:08Z6S1"
/translation="MNRIVFEDDAEVGSLIAAYLAKHDIVIEPRGRADRLILIT
PPAYLAKRLRLHLQSQTOQAKSIOESATLPKALRFGALITDPLNRAVQNGDFI
SUSTADFELMELATHAGQIMDRDLTLTLGVNTDGLDRSYDVAISRLRKKLLDSAA
EPRKRTIRNKGILFAPHAMETTG"
2108. 2440
/gene="STY1647"
/note="Pfam match to entry PF00072 response_reg, Response
regulator receiver domain, score 62.70, E-value 7.9e-15"
2582. 2794
/gene="STY1647"
/note="Pfam match to entry PF00486 trans_reg_C,
transcriptional regulatory protein, C terminal, score
62.90, E-value 1.8e-17"
complement(2901. 3260)
/gene="STY1648"
complement(2901. 3260)
/gene="STY1648"
/note="No significant database matches"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAD01893.1"
/db_xref="GI:16502737"
/db_xref="SPRMBL:08Z6S0"
/translation="MKORHPNLNIIISANANPSYMLINTKVTIONATYGEVYCYI

| Query Match | Best Local Similarity | Matches | Conservative | Mismatches | Indels | Gaps |
|---|--|---------|--------------|------------|--------|------|
| 0y 162 TCTTGGTGGAGCTATTTATGCTAAGTAACATTTACCCCGTGAAGTTAATGATCAATCA | 21.5%; Score 103.4; DB 1; Length 233050; | 148; | 0; | 56; | 1; | 1; |
| 0y 17897 TCTTCCGACCGCTTCCGACGAAAGTAACATTTACTCTG-TGTTTATGATAATACAA | 72.2%; Pred. No. 1.2e-16; | 148; | 0; | 56; | 1; | 1; |
| 0y 222 GAGNATTTGGGCTTATGAAATGCTCTTATTTGAATTAACAGTTGCGTCTTCTGCG | Score 103.4; DB 1; Length 233050; | 148; | 0; | 56; | 1; | 1; |
| 0y 17956 AGCAATATGCTCAATTAAGATGCTCTTATTTGAATTAACAGTTGCGTCTTCTGCG | Score 103.4; DB 1; Length 233050; | 148; | 0; | 56; | 1; | 1; |

| | | | |
|----|-------|--|-------|
| OY | 282 | TTTAGCATCTTCCTTGGCGTGCGAGTGCATGTGCAACTGACATCAGCCACTTGACAAGA | 341 |
| | | | |
| Dn | 18016 | TCTAGTCAATTCTTTTAGTAATTCGCCAACCAATATTGACAAATACCACACCTGAGCCCA | 18075 |
| OY | 342 | GTCGTCGGTACACACACAGTAAGAC | 366 |
| | | | |
| Dn | 18076 | CGGCTTCGCTTCACAGAGAGTAAGC | 18100 |

| | | | | | |
|------------|---|-------------|-----|--------|-----------------|
| RESULT 12 | ACOT9004/c | 134978 bp | DNA | Linear | HTG 24-AUG-2002 |
| LOCUS | ACOT9004/c | | | | |
| DEFINITION | Homo sapiens chromosome 15 clone RP11-137P11 map 15, *** SEQUENCING IN PROGRESS ***. 52 unordered pieces. | | | | |
| ACCESSION | AC079004 | | | | |
| VERSION | AC079004.2 | GI:11136775 | | | |
| KEYWORDS | HTG; HTGS_PHASE1. | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 134978) | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 2 (bases 1 to 134978) | | | | |
| AUTHORS | Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Batra,N., Bastien,V., Beda,F., Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Campatiano,A., Cattle,A., Choepel,T., Colangelo,M., Collins,S., Collumore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,I., Grand-pierre,N., Hagos,B., Heaford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Labocque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,C., Liu,G., Macdonald,P., Margulis,N., McCarthy,M., McEwan,P., McKernan,K., McNeeters,R., Meldrim,J., Menius,L., Mihova,T., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Sougnez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Ttigillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and zody.M. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (15-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA | | | | |
| REFERENCE | 3 (bases 1 to 134978) | | | | |
| AUTHORS | Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Boukhalter,B., Brown,A., Burkett,G., Campatiano,A., Cattle,A., Choepel,T., Colangelo,M., Collins,S., Collumore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,I., Grand-pierre,N., Hagos,B., Heaford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Labocque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,C., Liu,G., Macdonald,P., Margulis,N., McCarthy,M., McEwan,P., McKernan,K., McNeeters,R., Meldrim,J., Menius,L., Mihova,T., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Sougnez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Ttigillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and zody.M. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (32-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA | | | | |

COMMENT

On Nov 10, 2000 this sequence version replaced gi:3802827.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L10498

Center clone name: 137_P11

NOTE: This is a 'working draft' sequence. It currently consists of 52 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 164: contig of 164 bp in length
165 264: gap of 100 bp
265 848: contig of 584 bp in length
849 948: gap of 100 bp
949 1605: contig of 657 bp in length
1606 1705: gap of 100 bp
1706 2318: contig of 613 bp in length
2319 2418: gap of 100 bp
2419 3428: contig of 1010 bp in length
3429 3528: gap of 100 bp
3529 4191: contig of 663 bp in length
4192 4291: gap of 100 bp
4292 5090: contig of 799 bp in length
5091 5190: gap of 100 bp
5191 5963: contig of 773 bp in length
5964 6063: gap of 100 bp
6064 7369: contig of 1306 bp in length
7370 7469: gap of 100 bp
7470 8607: contig of 1138 bp in length
8608 8707: gap of 100 bp
8708 17424: contig of 8717 bp in length
17425 17524: gap of 100 bp
17525 18771: contig of 1247 bp in length
18772 18871: gap of 100 bp
18872 19924: contig of 1053 bp in length
19925 20024: gap of 100 bp
20025 21370: contig of 1346 bp in length
21371 21470: gap of 100 bp
21471 23256: contig of 1786 bp in length
23257 23356: gap of 100 bp
23357 25010: contig of 1654 bp in length
25011 25110: gap of 100 bp
25111 26876: contig of 1766 bp in length
26877 26976: gap of 100 bp
26977 28146: contig of 1170 bp in length
28147 28246: gap of 100 bp
28247 29546: contig of 1300 bp in length
29547 29646: gap of 100 bp
29647 31979: contig of 2333 bp in length
31980 32079: gap of 100 bp
32080 33806: contig of 1727 bp in length
33807 33906: gap of 100 bp
33907 35287: contig of 1381 bp in length
35288 35387: gap of 100 bp
35388 36754: contig of 1367 bp in length
36755 36854: gap of 100 bp
36855 38157: contig of 1303 bp in length
38158 38257: gap of 100 bp
38258 40144: contig of 1887 bp in length
40145 40244: gap of 100 bp
40245 42144: contig of 1900 bp in length
42145 42244: gap of 100 bp

42245 45014: contig of 2770 bp in length
45015 45114: gap of 100 bp
45115 47299: contig of 2185 bp in length
47300 47399: gap of 100 bp
47400 49714: contig of 2315 bp in length
49715 49814: gap of 100 bp
49815 51867: contig of 2053 bp in length
51868 51967: gap of 100 bp
51968 54711: contig of 2744 bp in length
54712 54811: gap of 100 bp
54812 57606: contig of 2795 bp in length
57607 57706: gap of 100 bp
57707 60300: contig of 2594 bp in length
60301 60400: gap of 100 bp
60401 62959: contig of 2559 bp in length
62960 63059: gap of 100 bp
63060 64544: contig of 1485 bp in length
64545 64644: gap of 100 bp
64645 66943: contig of 2289 bp in length
66944 67043: gap of 100 bp
67044 68675: contig of 1632 bp in length
68676 68775: gap of 100 bp
68776 71620: contig of 2845 bp in length
71621 71720: gap of 100 bp
71721 74270: contig of 2550 bp in length
74271 74370: gap of 100 bp
74371 78405: contig of 4035 bp in length
78406 78505: gap of 100 bp
78506 81246: contig of 2741 bp in length
81247 81346: gap of 100 bp
81347 83767: contig of 2421 bp in length
83768 83867: gap of 100 bp
83868 86612: contig of 2745 bp in length
86613 86712: gap of 100 bp
86713 89372: contig of 2660 bp in length
89373 89472: gap of 100 bp
89473 93415: contig of 3943 bp in length
93416 93515: gap of 100 bp
93516 98713: contig of 5198 bp in length
98714 98813: gap of 100 bp
98814 103055: contig of 4242 bp in length
103056 103155: gap of 100 bp
103156 109896: contig of 6641 bp in length
109897 110096: gap of 100 bp
110097 117716: contig of 7620 bp in length
117717 117816: gap of 100 bp
117817 124867: contig of 7051 bp in length
124868 124967: gap of 100 bp
124968 133690: contig of 8723 bp in length
133691 133790: gap of 100 bp
133791 134978: contig of 1188 bp in length.

FEATURES

source

1..134978

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-137P11"

/clone_lib="RPC1-11 Human Male BAC"

1..164

/note="assembly-fragment"

clone_end:SP6

vector_side:left"

265..848

/note="assembly-fragment"

misc-feature

Query Match

8.8%; Score 42.4; DB 2; Length 134978;

Best Local Similarity 53.7%; Pred. No. 1.2; Mismatches 88; Conservative 0; Indels 76; Gaps 0;

DB 3207 ATTGACTCTGCGCTCTCCGCGCCTCATTCCTTATCCCTCTTCCTTACCTATG 3148


```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
*      1      2073: contig of 2073 bp in length
*
* 2074 2173: gap of 100 bp
*
* 2174 3577: contig of 1404 bp in length
*
* 3578 3677: gap of 100 bp
*
* 3678 4992: contig of 1315 bp in length
*
* 4993 5092: gap of 100 bp
*
* 5093 7181: contig of 2089 bp in length
*
* 7182 7281: gap of 100 bp
*
* 7282 9225: contig of 1944 bp in length
*
* 9226 9325: gap of 100 bp
*
* 9326 11777: contig of 2452 bp in length
*
* 11778 11877: gap of 100 bp
*
* 11878 14570: contig of 2693 bp in length
*
* 14571 14670: gap of 100 bp
*
* 14671 17522: contig of 2852 bp in length
*
* 17523 17622: gap of 100 bp
*
* 17623 20980: contig of 3358 bp in length
*
* 20981 21080: gap of 100 bp
*
* 21081 25542: contig of 4462 bp in length

```

```

FEATURES
    * 154051 164184: contig of 10134 bp in length
      location/Qualifiers
        source 1. 164184
              /organism="Homo sapiens"

```

```
misc_feature      /map="8"  
                  /clone="PPI-111G3"  
                  /clone_lib="Rpci-11 Human Male BAC"  
1..2073  
/note="assembly-fragment  
clone_end:SP6  
vector_slide:left"  
2174..3577  
/note="assembly-fragment"  
3678..4992  
/note="assembly-fragment"  
5093..7181  
/note="assembly-fragment"  
7282..9225  
/note="assembly-fragment"  
9226..  
/note="assembly-fragment"  
9326..11177  
/note="assembly-fragment"  
11878..14570  
/note="assembly-fragment"  
14671..17522
```


| | | |
|---------------|---------------------------|----------------------------|
| repeat_region | /clone="CTD-250769" | |
| | 40. .165 | /clone_1lb="CTD2 Human BAC |
| repeat_region | /rpt_family="FLAM_C" | |
| | complement(216. .670) | |
| repeat_region | /rpt_family="LTR32" | |
| | 883. .1192 | |
| repeat_region | /rpt_family="AluSx" | |
| | 2230. .2449 | |
| repeat_region | /rpt_family="L1MC4a" | |
| | complement(2471. .2688) | |
| repeat_region | /rpt_family="MER51A" | |
| | complement(21717. .2970) | |
| repeat_region | /rpt_family="MER51A" | |
| | complement(3047. .3095) | |
| repeat_region | /rpt_family="MER51A" | |
| | 3132. .4333 | |
| repeat_region | /rpt_family="L1MC4a" | |
| | complement(4363. .4956) | |
| repeat_region | /rpt_family="MER50" | |
| | complement(4957. .5262) | |
| repeat_region | /rpt_family="AluY" | |
| | complement(5263. .5330) | |
| repeat_region | /rpt_family="MER50" | |
| | 5337. .6142 | |
| repeat_region | /rpt_family="L1MC4a" | |
| | 6143. .6441 | |
| repeat_region | /rpt_family="AluY" | |
| | 6442. .6513 | |
| repeat_region | /rpt_family="L1MC4a" | |
| | 6519. .6602 | |
| repeat_region | /rpt_family="TA)n" | |
| | 6604. .6758 | |
| repeat_region | /rpt_family="LTR8A" | |
| | 6759. .6795 | |
| repeat_region | /rpt_family="MER1B" | |
| | 6956. .7078 | |
| repeat_region | /rpt_family="L1MC4" | |
| | complement(7702. .8619) | |
| repeat_region | /rpt_family="L1P1A16" | |
| | complement(9197. .9374) | |
| repeat_region | /rpt_family="L1M" | |
| | complement(10565. .10788) | |
| repeat_region | /rpt_family="MIR" | |
| | complement(10942. .11305) | |
| repeat_region | /rpt_family="t2" | |
| | 11691. .11880 | |
| repeat_region | /rpt_family="t2" | |
| | 11892. .11924 | |
| repeat_region | /rpt_family="(TGAA)n" | |
| | 12948. .12972 | |
| repeat_region | /rpt_family="AT-rich" | |
| | 13462. .13752 | |
| repeat_region | /rpt_family="AluSx" | |
| | complement(13989. .14140) | |
| repeat_region | /rpt_family="MIR" | |
| | 14464. .15015 | |
| repeat_region | /rpt_family="MER57A" | |
| | 15114. .15682 | |
| repeat_region | /rpt_family="MIR" | |

```

repeat_region      15685. .21818  
/rpt_family="LIPa3"  
repeat_region      22008. .22232  
/rpt_family="MIR"  
repeat_region      22233. .22378  
/rpt_family="(ATG)n"  
repeat_region      complement(23783. .24116)  
/rpt_family="MER1b"  
repeat_region      complement(24181. .24295)  
/rpt_family="MIR"  
repeat_region      24650. .24759  
/rpt_family="MIR"  
repeat_region      24765. .24812  
/rpt_family="(CA)n"  
repeat_region      25178. .25219  
/rpt_family="AT-rich"  
repeat_region      25318. .25403  
/rpt_family="AT-rich"  
repeat_region      26448. .26526  
/rpt_family="LIM4"  
repeat_region      complement(26824. .27132)  
/rpt_family="AluY"  
repeat_region      27637. .27776  
/rpt_family="L2"  
repeat_region      complement(28202. .28550)  
/rpt_family="MSTD"  
repeat_region      complement(28805. .29162)  
/rpt_family="THE1c"  
repeat_region      29912. .29965  
/rpt_family="Ct-rich"

```

| | | | | |
|-----------------------|-----------------|----------------|-----------|----------------|
| Query Match | 8.88; | Score 42.4; | DB 9; | Length 167577; |
| Best Local Similarity | 53.78; | Pred. No. 1.2; | | |
| Matches 88; | Conservative 0; | Mismatches 76; | Indels 0; | Gaps 0; |

26 ATTGACACTATCCTTACCCACGCTCAACAGTTAATAACCTGCCAGCAATAAGGATC 85

Db 65598 ATTGACTCTGGCTCCTTCCCTGGCCTCATTCCTATATCCCTCTTCCGTGTCCTTCACTATG 656

86 TTGTTAACTTAAGTCAAAAAAATAGCGAATTTCCACGACAAAGCTAAATATCGCA 145

DD 03030 TGGTATACCTTGAATCAATCTAATAATACATGTAAAGCAATACAGAAAAA

[illegible]

RESULT 15
*E0113C

| LOCUS | AL591126 | 190015 bp | DNA | linear | HTG 29-JUN-93 |
|------------|---|-----------|-----|--------|---------------|
| DEFINITION | Mus musculus chromosome 11 clone PD3-188A3 *** SEQUENCING INFORMATION *** | | | | |

PROGRESS ***, 2 unordered pieces.
ACCESSION AI591126

```

VERSION      01.2100551,
KEYWORDS     HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.

```

ORGANISM Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS
Phillimore, B.

JOURNAL
Submitted (28-JUN-2002) Wellcome Trust Sanger Institute, Hinxton
Cambridgeshire CB10 1SA UK E-mail: aravind.srinivasan@sanger.ac.uk

COMMENT On Jul 2, 2002 this sequence version replaced at: 21655322

Center: Wellcome Trust Sanger Institute

Web site: <http://www.sanger.ac.uk>

----- Project Information -----

```

Center project name: bm188a3
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator; 6% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 93% of reads
Consensus quality: 189901 bases at least Q40
Consensus quality: 189915 bases at least Q30
Consensus quality: 189915 bases at least Q20
Insert size: 189915; sum-of-contigs
Insert size: 188814; 4.0% error; agarose-fp
Quality coverage: 18.04x in Q20 bases; sum-of-contigs Quality
coverage: 20.10x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 11148: contig of 11148 bp in length
* 11149 11248: gap of 100 bp
* 111249 190015: contig of 78767 bp in length.
Location/Qualifiers
1.190015
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-188A3"
/clone_11b="RPCT-23"
1.11148
/note="assembly_fragment:07880.0"
111249.190015
/note="assembly_fragment:07880.1"
BASE COUNT 60196 a 40379 c 37523 g 51817 t 100 others
ORIGIN
Query Match 8.8% Score 42.2; DB 2; Length 190015;
Best Local Similarity 58.3%; Pred. No. 1.3;
Matches 74; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

OY 103 AAAAATAGCAATTTTCACAGCAAAAGCTAAATATGCAAAAACCTCAGTAAAT 162
||||| ||| || ||||| ||||| ||||| ||||| |||||
DB 21377 AAAAATAGCAATTTTCACAGCAAAAGCTAAATATGCTACATATATATATAT 21436
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 163 CTTCGTCGAGCTATTATGCTAAGTAACTTTACCCCTGAAGTTAATGATCAATCAAG 222
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 21437 ATTGTATTATGTTATTTACATGACATTAACCATCTTTTCYAAAGAGAAATATAT 21496
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 223 AGAGATG 229
|||||
DB 21497 AGAGCTG 21503
|||||

```

Search completed: February 7, 2003, 20:36:34
Job time : 2920 secs


```

FT      mutation      replace (235..237, TAG)
FT      /*tag= e
FT      /label= "Amber-mutation_V20
FT      /note= "An amber mutation at this position does not
FT      abolish beta-lactam resistance"
FT      CDS
FT      239..364
FT      /*tag= f
FT      /product= "BLR protein"
FT      /replace (247..249, TAG)
FT      /*tag= g
FT      /label= "Amber-mutation_L24
FT      /note= "An amber mutation at this position abolishes
FT      beta-lactam resistance, but this can be reversed
FT      using an amber suppressor"
FT      mutation      replace (292..294, TAG)
FT      /*tag= h
FT      /label= "Amber-mutation_L39
FT      /note= "An amber mutation at this position abolishes
FT      beta-lactam resistance, but this can be reversed
FT      using an amber suppressor"
FT      primer_bind    332..335
FT      /*tag= 1
FT      /note= "Primer R3 (not given in the specification)"
FT      338..339
FT      /*tag= 1
FT      /note= "Insertion point for mini-TnphoA transposon in
FT      E. coli strain RM583"
FT      primer_bind    448..465
FT      /*tag= k
FT      /note= "Primer R1 (not given in the specification)"
FT      470..481
FT      /*tag= 1
FT      /label= "ORE71-promoter
FT      /note= "Promoter for ORE71 (putative histone-like
FT      negative regulator gene"
FT      WO200177176-A2.
FT      18-OCT-2001.
FT      PD
FT      18-OCT-2001.
FT      PF
FT      06-APR-2001; 2001WO-US11363.
FT      XX
FT      06-APR-2000; 2000US-195505P.
FT      PR
FT      14-JUL-2000; 2000US-218380P.
FT      XX
FT      (TUFT ) TUFTS COLLEGE.
FT      PA
FT      Levy SB, McMurry LM;
FT      PI
FT      MPI: 2002-010898/01.
FT      XX
FT      P-PSDB; AAM52607.
FT      DR
FT      XX
FT      New membrane protein, designated Beta Lactam-358 polypeptides, that
FT      affect susceptibility to antibiotics which affect peptidoglycan
PT      synthesis in microbes, useful for identifying modulators for treating
PT      infections -
PT      Claim 1, Fig 1A-B, 104pp; English.
XX      PS
XX      This sequence represents a beta-lactam resistance (blr) gene from
CC      Escherichia coli. The blr gene encodes a 41 amino acid transmembrane
CC      protein (also designated Beta Lactam-358 in the specification) with a
CC      molecular weight of 4556 Daltons which confers resistance to beta-lactam
CC      antibiotics. The blr gene was initially identified in a 602 bp
CC      antidiotic. The blr gene is a member of an antibiotic-susceptible
CC      hypohelical intergenic region of the genome of an antibiotic-resistant
CC      strain of E. coli, RM583. In the RM583 strain, the blr gene is disrupted
CC      by insertion of the transposon mini-TnphoA. A clone bearing an ORF (open
CC      reading frame)-containing 358 bp segment of the "intergenic" region was
CC      found to restore beta-lactam resistance in RM358. Certain amber
CC      mutations introduced into the 358 bp clone were found to abolish
CC      beta-lactam resistance and resistance could be restored using an
CC      amber suppressor, indicating that a protein encoded by the clone was
CC      responsible for mediating beta-lactam resistance. It is suggested that

```

```

CC      the BLR protein may act by increasing a beta-lactamase activity in a
CC      cell, possibly being a component of an uncharacterised membrane-bound
CC      efflux pump capable of expelling beta-lactams from the periplasm. The
CC      BLR protein may be used in methods to identify compounds that modulate
CC      antibiotic resistance in a bacterium. The BLR protein, gene, BLR protein
CC      homologues, BLR agonists or antagonists, and anti-BLR antibodies may be
CC      used to treat bacterial infection, particularly by an organism resistant
CC      to antibiotics that affect peptidoglycan synthesis. They may also be
CC      used in screening assays, in vaccines, and in diagnostic assays. BLR
CC      nucleotides are useful in gene therapy applications to modulate BLR
CC      expression, and to detect BLR expression in a biological sample.
XX      SQ
XX      Sequence 481 BP; 139 A; 95 C; 100 G; 147 T; 0 other:
XX
XX      Query Match      100.0%; Score 481; DB 24; Length 481;
XX      Best Local Similarity 100.0%; Pred. No. 1.7e-129;
XX      Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1      TGCCCTCATGAGTATGATGCTATTTGACACTATCCTTTACCCAGGCTCAACAGTTT 60
DB      1      TGCCTCTCATTTGAAGTATGATGCTATTTGACACTATCCTTTACCCAGGCTCAACAGTTT 60
QY      61      AATAACCTGCCAGCAATPAAGGATGTTTAACTTAAGTCATAAAATAAGCGAATTTTC 120
DB      61      AATAACCTGCCAGCAATPAAGGATGTTTAACTTAAGTCATAAAATAAGCGAATTTTC 120
QY      121      CAAGCAACAAAGCTAAATATTCGCAAAACCTCAGTAAATCTTGCTGAGCTATTATT 180
DB      121      CAAGCAACAAAGCTAAATATTCGCAAAACCTCAGTAAATCTTGCTGAGCTATTATT 180
QY      181      GCTAAGTAACATTTAACCCCGTAAGTTAATGATCAATCAAGACAGATGCGCTGTAAT 240
DB      181      GCTAAGTAACATTTAACCCCGTAAGTTAATGATCAATCAAGACAGATGCGCTGTAAT 240
QY      241      GATAGCTCTTATTTGAATTAACAGGTTGATGCTTTCGTTTCAGTCAATCTTCTTGG 300
DB      241      GATAGCTCTTATTTGAATTAACAGGTTGATGCTTTCGTTTCAGTCAATCTTCTTGG 300
QY      301      CGTGGCAGTGCATTTGACCAATATCAGCCACCTGTAACAGAGTCTTGCTGTAACACAA 360
DB      301      CGTGGCAGTGCATTTGACCAATATCAGCCACCTGTAACAGAGTCTTGCTGTAACACAA 360
QY      361      GTAGCTCTGCACCTTGTGAGCGACATCTGCCGCTCGGGTGCATGTTTTCATTGTCG 420
DB      361      GTAGCTCTGCACCTTGTGAGCGACATCTGCCGCTCGGGTGCATGTTTTCATTGTCG 420
QY      421      GATATTAACCGAATTTATTTATCTTGTTGATGTTTGTGTTGATGTTGATGATGATG 480
DB      421      GATATTAACCGAATTTATTTATCTTGTTGATGTTTGTGTTGATGTTGATGATGATG 480
QY      481      A 481
DB      481      A 481
XX
XX      RESULT 2
XX      ABA90521/C
XX      ID      ABA90521 standard; DNA; 2365589 BP.
XX      AC      ABA90521;
XX      DT      16-MAY-2002 (first entry)
XX      DE      Genomic sequence of Lactococcus lactis IL1403.
XX      KW      Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds-
XX      OS      Lactococcus lactis IL1403.
XX      PN      FR2807446-A1.
XX      PD      12-OCT-2001.
XX

```


KW antiInflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX Homo sapiens.
OS
PM WO200200928-A2.
XX
XX 03-JAN-2002.
PD
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
PR
XX (EPIC-) EPIGENOMICS AG.
PA
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-130909/17.
DR
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX
PS Claim 1; SEQ ID NO 341; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/nutritional bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 9997 BP; 2290 A; 91 C; 2181 G; 5435 T; 0 other;
XX
Query Match 8.1%; Score 38.8; DB 24; Length 9997;
Best Local Similarity 54.1%; Pred. No. 0.77;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
XX
QY 13 AAGTATGATGGCTATTGTGACACTATCTTACCCGCTCAACAGTTTATTAACCTGCCA 72
DB 5923 AATATATTATACCACTAATACCAACAAACACACACTTAATCTTTACATTAATAACATA 5864
QY 73 GCAATTAAGGAGTGTGTTTAAGTCAAAAAAATAGCGAATTTTCCACGACACAAAA 132
DB 5863 TATATTTCCACTATTTATTAATAAATTACTCAAAAACACAAATTTCTCCAACACCAAAA 5804
QY 133 GCTAAATATGCGAAAAAACCTCAGTAA 158
DB 5803 AAAAAATTAACCTAATAAAAAAATAA 5778
XX
RESULT 5
ABL34484/C
ID ABL34484 standard; DNA; 9997 BP.
XX
AC ABL34484;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human metastasis associated gene SEQ ID NO: 37.
XX
XX Metastasis associated gene; cytosinatic; gene therapy; cancer;
KW
XX cytosine methylation; gene; ds.
XX
XX Homo sapiens.
OS
XX WO200177376-A2.
PN
XX

PD 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-EP03970.
PF
XX
XX 06-APR-2000; 2000DE-1019058.
PR
XX 07-APR-2000; 2000DE-1019173.
PR
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
PA
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-010922/01.
DR
XX
XX New nucleic acid derived from chemically treated metastasis genes,
PT useful for diagnosis of cancers by analysis of cytosine methylation,
PT also for treatment -
XX
XX Claim 1; SEQ ID NO 37; 23pp + Sequence Listing; English.
XX
XX The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention.
XX
SQ Sequence 9997 BP; 2290 A; 91 C; 2181 G; 5435 T; 0 other;
XX
Query Match 8.1%; Score 38.8; DB 24; Length 9997;
Best Local Similarity 54.1%; Pred. No. 0.77;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
XX
QY 13 AAGTATGATGGCTATTGTGACACTATCTTACCCGCTCAACAGTTTATTAACCTGCCA 72
DB 5923 AATATATTATACCACTAATACCAACAAACACACACTTAATCTTTACATTAATAACATA 5864
QY 73 GCAATTAAGGAGTGTGTTTAAGTCAAAAAAATAGCGAATTTTCCACGACACAAAA 132
DB 5863 TATATTTCCACTATTTATTAATAAATTACTCAAAAACACAAATTTCTCCAACACCAAAA 5804
QY 133 GCTAAATATGCGAAAAAACCTCAGTAA 158
DB 5803 AAAAAATTAACCTAATAAAAAAATAA 5778
XX
RESULT 6
ABK31511/C
ID ABK31511 standard; DNA; 47108 BP.
XX
XX ABK31511;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #177.
XX
XX Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytosinatic; mutant; ds.
XX
XX Homo sapiens.
OS
XX Synthetic.
OS
XX WO200200926-A2.
PN
XX 03-JAN-2002.
PD
XX 29-JUN-2001; 2001WO-EP07472.
PF
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
PR
XX
XX (EPIC-) EPIGENOMICS AG.
PA
XX

PI Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2002-147896/19.
XX
XX Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction -
XX
XX
XX Claim 1; SEQ ID NO 354; 24pp; English.
XX
XX The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
CC for detecting the cytosine methylation state (CpG islands) of these
CC genes, and a method for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with signal transduction.
CC The genomic DNA can be obtained from cells or cellular components which
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
CC histologic object slides, and all their possible combinations. The
CC sequences of the invention are useful for the diagnosis and therapy of
CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK3158-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
XX Sequence 47108 BP: 15349 A; 429 C; 8736 G; 22594 T; 0 other;
SQ
Query Match 8.0%; Score 38.4; DB 24; Length 47108;
Best Local Similarity 48.2%; Pred. No. 1.8;
Matches 108; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
OY 43 ACCACGCTCAGACTTAACTGCGCAGCAATAGGATGCTGTTAACTTAATCA 102
DB 19108 ACACCCACACCAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 19049
OY 103 AAAAAATAGCGAATTTCCACAGCAACAAAGTAATATGCAAAAACTCAGTAAAT 162
DB 19048 TAAAAACAAAAAAGTAAATATGCAAAAACTTAACTTAACTTAACTTAACT 16989
OY 163 CTGCTGAGCTTATTTGCTAAGTAACATTTACCCCTGAGTTAATGATCAATCAAG 222
DB 18988 AATATATTAATATTAACCTTAATCAACAATTAACCTTAATTTTAAATTTAA 18929
OY 223 AGAGATGTGGGCTGTATGAATGCTTATTAATTAACAGGTT 266
DB 18928 ACCCAATTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 18885

RESULT 7
ABL70143/C
ID ABL70143 standard; DNA: 15881 BP.
XX
XX ABL70143;
XX
XX 01-JUL-2002 (first entry)
XX
XX Chemically treated cell signalling DNA sequence#17.
XX
XX Cell signalling; cytosine methylation; cell signalling disease;
KW cancer; tumour; cytostatic; ds.
XX
XX unidentified.
OS
XX
XX WO200202807-A2.
XX
XX PD 10-JAN-2002.
XX

PF 29-JUN-2001; 2001WO-EP07471.
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2002-154758/20.
XX
XX Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signalling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signalling -
XX
XX
XX Claim 1; SEQ ID NO 33; 24pp+sequence listing; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
XX Sequence 15881 BP: 4515 A; 204 C; 3588 G; 7574 T; 0 other;
SQ
Query Match 7.9%; Score 37.8; DB 24; Length 15881;
Best Local Similarity 52.2%; Pred. No. 1.8;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
OY 102 AAAAAATAGCGAATTTCCACAGCAACAAAGTAATATGCAAAAACTCAGTAAATA 161
DB 14565 AAAAAAATCAAAAAACACACAAAAAAGTAAATATGCAAAAAACACACAC 14506
OY 162 TCTTGTGAGCTTATTTGCTAAGTAACATTTACCCCTGAGTTAATGATCAATCAA 221
DB 14505 CCTCCATTCATAATTTATTAATTAATTAATTAATTTTCTCTTAATTAATTAATTA 14446
OY 222 GAGAGATGTGGGCTGTATGAATGCTTATTAATTAAC 262
DB 14445 AATATATCTATTTTAAATTAACCTCTTAATTAATTAATTAATTAATTAATTA 14405

RESULT 8
ABL32260/C
ID ABL32260 standard; DNA: 15881 BP.
XX
XX ABL32260;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 233.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; neoplastic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
OS
XX
XX Homo sapiens.
XX

XX W0200200928-A2.
XX
XX 03-JAN-2002.
PD
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX 30-JUN-2000; 2000DE-1032529.
PR
PR 01-SEP-2000; 2000DE-1043626.
XX
XX (EPIC-) EPIGENOMICS AG.
PA
PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
DR
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX Claim 1; SEQ ID NO 233; 32pp + Sequence Listing; German.
PS
XX
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 15881 BP; 4515 A; 204 C; 3588 G; 7574 T; 0 other;

Query Match 7.9%; Score 37.8; DB 24; Length 15881;
Best Local Similarity 52.2%; Pred. No. 1.8;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 102 AAAAAAATGCGAATTTTCCAGACAAAGCTAAATATCGAAAAAACCCTCAGTAAAA 161
DB 14565 AAAAAAATGCGAATTTTCCAGACAAAGCTAAATATCGAAAAAACCCTCAGTAAAA 14506
QY 162 TCTTGGTGGAGCTATTATGCTAAGTACATTTACCCCTGAAGTTAATGATCAATCA 221
DB 14505 CCTCCATTCGAATTTAATTAATATAAAATTTTCTCCTAATATAAAAAATATAAA 14446
QY 222 GAGAGATGTGGCTGTAATGAATCGTCTTATTTGAATTAACA 262
DB 14445 AATATATCTATTTTAAATAAACCTCTATACCTTTAATA 14405

RESULT 9
AAS61069/c
ID AAS61069 standard; DNA; 15881 BP.
XX
XX AAS61069;
AC
XX
DT 29-JAN-2002 (first entry)
XX
XX Human gene regulation-associated gene oligonucleotide #24.
DE
XX
XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
KW cardiac damage; inflammatory response; haemophilia; Werner syndrome;
KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
KW renal disease; preclampsia; cardiac allograft vascular disease;
KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
KW immunostimulant; cardiac; anti-inflammatory; coagulant; antiplatelet;
KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cystostatic.
XX
XX Homo sapiens.
OS
XX
XX W0200177375-A2.
PN
XX
XX 18-OCT-2001.
PD

XX
XX 06-APR-2001; 2001WO-EP03368.
PF
XX
XX 06-APR-2000; 2000DE-1019058.
PR
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043626.
XX
XX (EPIC-) EPIGENOMICS AG.
PA
PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-017470/02.
DR
XX
XX New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease -
XX
XX Claim 1; SEQ ID NO 25; 26pp; English.
PS
XX
XX The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC preclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 15881 BP; 4515 A; 204 C; 3588 G; 7574 T; 0 other;

Query Match 7.9%; Score 37.8; DB 24; Length 15881;
Best Local Similarity 52.2%; Pred. No. 1.8;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 102 AAAAAAATGCGAATTTTCCAGACAAAGCTAAATATCGAAAAAACCCTCAGTAAAA 161
DB 14565 AAAAAAATGCGAATTTTCCAGACAAAGCTAAATATCGAAAAAACCCTCAGTAAAA 14506
QY 162 TCTTGGTGGAGCTATTATGCTAAGTACATTTACCCCTGAAGTTAATGATCAATCA 221
DB 14505 CCTCCATTCGAATTTAATTAATATAAAATTTTCTCCTAATATAAAAAATATAAA 14446
QY 222 GAGAGATGTGGCTGTAATGAATCGTCTTATTTGAATTAACA 262
DB 14445 AATATATCTATTTTAAATAAACCTCTATACCTTTAATA 14405

RESULT 10
ABK31182/c
ID ABK31182 standard; DNA; 15881 BP.
XX
XX ABK31182;
AC
XX
DT 23-APR-2002 (first entry)
XX
XX Signal transduction associated gene modified DNA #13.
DE
XX
XX Human; signal transduction associated gene; cytosine methylation state;
KW

KW CPG island; signal transduction associated disease; solid tumour; cancer;
KM antitumour; cytostatic; mutant; ds.
OS Homo sapiens.
OS Synthetic.
PN WO200200926-A2.
XX
XX
PD 03-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-EP07472.
PF
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX Olek A, Plepenbrock C, Berlin K;
PI
XX WPI: 2002-147896/19.
DR
XX
XX Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction -
XX
XX
PS Claim 1; SEQ ID No 25; 24pp; English.
XX
XX The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
CC for detecting the cytosine methylation state (CPG islands) of these
CC genes, and a method for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with signal transduction.
CC The genomic DNA can be obtained from cells or cellular components which
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
CC histologic object slides, and all their possible combinations. The
CC sequences of the invention are useful for the diagnosis and therapy of
CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
CC
XX
SQ Sequence 15881 BP; 4515 A; 204 C; 3588 G; 7574 T; 0 other;
SO
Query Match 7.9%; Score 37.8; DB 24; Length 15881;
Best Local Similarity 52.2%; Pred. No. 1.8;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 102 AAAAAATAGCGAATTTCCACAGCAAAAGCTAATATCGAAAAAACCCTCAGTAAAA 161
DB 14565 AAAAAATAGCGAATTTCCACAGCAAAAGCTAATATCGAAAAAACCCTCAGTAAAA 161
QY 162 TCTTGGTGGAGCTATTATTTGCTAAGTACATTTACCCCTGAAGTTAATGATCAATCAA 221
DB 14505 CCGCCATTCGAATTTAATTAATAATAAATTTTCCCTAATTAATAAATAAATAA 14446
QY 222 GAGAGATGTGGGCTGTAATGAATCGTCTTATTTGAATTAACA 262
DB 14445 AATATATCTATTTTAAATAAATCCTTATCTTATA 14405

RESULT 11
AAC85009
ID AAC85009 standard; DNA; 15456 BP.
XX
AC AAC85009;
XX

DT 08-MAY-2001 (first entry)
XX
XX Nucleotide sequence of bovine PIV3 SF strain.
DE
XX
XX Parainfluenza virus; PIV; infectious; human; nucleocapsid protein; BPIV;
KM nucleocapsid phosphoprotein; polymerase protein; human; bovine; HPIV;
KM chimeric; vaccine; immune response; HPIV1; HPIV2; HPIV3; BPIV3; ds.
XX
XX Bovine parainfluenza virus.
XX
XX WO200104320-A1.
XX
XX 18-JAN-2001.
XX
XX 16-JUN-2000; 2000WO-US17066.
PF
XX
XX 09-JUL-1999; 99US-0143134.
PR
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Schmidt AC, Skiadopoulos MH, Collins PL, Murphy BR, Bailly JE;
PI Durbin AP;
XX
XX WPI: 2001-081053/09.
DR
XX
XX Isolated human-bovine chimeric parainfluenza virus (PIV), useful in an
PT attenuated vaccine to elicits an immune response against one or more
PT virus(es) selected from human PIV1 (HPIV1), HPIV2 and HPIV3 -
XX
XX
PS Disclosure; Fig 7A-G; 148pp; English.
XX
XX The invention relates to an isolated human-bovine chimeric parainfluenza
CC virus (PIV) that is infectious and attenuated in humans. The virus
CC comprises a major nucleocapsid protein, a nucleocapsid phosphoprotein,
CC a large polymerase protein, and a partial or complete PIV background
CC genome, or antigenome of a human PIV (HPIV) or bovine PIV (BPIV),
CC combined with one or more heterologous gene(s) or genome segment(s) of a
CC different PIV to form a human-bovine chimeric PIV genome or antigenome.
CC The chimeric PIV is useful in attenuated vaccines to elicit immune
CC response against one or more virus(es) selected from HPIV1, HPIV2 and
CC HPIV3 and most preferably against HPIV3. The present sequence represents
CC the nucleotide sequence of bovine PIV3 SF strain.
XX
XX
SQ Sequence 15456 BP; 5872 A; 2656 C; 2935 G; 3993 T; 0 other;
SO
Query Match 7.7%; Score 37.2; DB 22; Length 15456;
Best Local Similarity 54.3%; Pred. No. 2.6;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 208 AATGATCAATCAAGAGATGTGGCTGTAATGAATCGTCTTATTTGAATTAACAGGTTG 267
DB 3926 AATGGAAGGTCAAAAAGACAGATGAGCATTAAGTATCTAGATGATCAAGTAA 3985
QY 268 GATCGTTTGTGCGTTTCAGTATCTTCTGGCGGAGAGACATGACAAATGATCA 327
DB 3986 CAAAGTTTGTGCTGTGATCATTTGCCACTTGCGTTGGCTAGTATCACTGGAATGATCA 4045
QY 328 GCCACCTGAACAGAGTGC 345
DB 4046 GGAACCTCAACAGGCTGC 4063

RESULT 12
AAD27766
ID AAD27766 standard; DNA; 15456 BP.
XX
AC AAD27766;
XX
XX 18-APR-2002 (first entry)
DT
XX Bovine parainfluenza virus 3 (BPIV3) SF DNA.
DE
XX Parainfluenza virus; PIV; nucleocapsid protein; polymerase protein;
XX

KW HIV; human PIV, nucleocapsid phosphoprotein; bovine PIV3; BPIV3;
 KW prophylaxis; infection; vaccine; immune response; immunostimulant;
 KW virucide; SF DNA; ds.
 XX
 OS Bovine parainfluenza virus.
 XX
 PN WO200202605-A2.
 XX
 PD 10-JAN-2002.
 XX
 XX 05-JUL-2001; 2001WO-US21527.
 PF
 XX 05-JUL-2000; 2000US-215809P.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Skladopoulos MH, Collins PL, Murphy BR, Schmidt AC;
 PI WPI: 2002-140084/18.
 XX
 DR Isolated human-bovine chimeric parainfluenza virus (PIV), useful in an
 PT attenuated vaccine to elicits an immune response against one or more
 PT virus(es) selected from human PIV1 (HPIV1), HPIV2 and HPIV3 -
 XX
 PS Example 1; Fig 2; 154pp; English.
 XX
 CC The present invention relates to chimeric human-bovine parainfluenza
 CC viruses (PIV) that are infectious and attenuated in humans and other
 CC mammals. The viruses comprise a major nucleocapsid protein (N), a
 CC nucleocapsid phosphoprotein (P), a large polymerase protein (L) and
 CC a partial or complete PIV background genome or antigenome of a human
 CC PIV (HPIV) or bovine PIV (BPIV), combined with one or more heterologous
 CC genes or genome segments of a N, P, L, or M gene of a different PIV to
 CC form a human-bovine chimeric PIV genome or antigenome. The invention
 CC also relates to methods and compositions incorporating human-bovine
 CC chimeric PIV for prophylaxis and treatment of PIV infection. The
 CC chimeric PIVs are useful in attenuated vaccines to elicit an immune
 CC response against one or more viruses selected from HPIV1, HPIV2 and
 CC HPIV3. The present sequence is bovine parainfluenza virus 3 (BPIV3)
 CC SF DNA. This sequence is used in the exemplification of the invention.
 XX
 SQ Sequence 15456 BP; 5871 A; 2653 C; 2936 G; 3996 T; 0 other;
 XX
 Query Match 7.7%; Score 37.2; DB 24; Length 15456;
 Best Local Similarity 54.3%; Pred. No. 2.6;
 Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
 XX
 QY 208 AATGATCATCATAGAGAGATGGGCTGTAATGATCGTCTTATGATTAACAGGTG 267
 DB 3926 AATGGAAGGTCAAAAGACAGGTATGAGAGCATTAAGTATGATGATCAAGTTA 3985
 QY 268 GATCGTTCTTCTGCTTCAGTCAATCTTGTGGCGGAGTACATTGACAACTATCA 327
 DB 3986 CAAGGTTTGGCTCTGATCATTTGCTGCTTGGCTGAGTATACAGGAATATATCA 4045
 QY 328 GCCACCTGAACAGAGTGC 345
 DB 4046 GGAACCTCTACAGGCTGC 4063
 XX
 RESULT 13
 ID ABL34168/c
 XX ABL34168 standard; DNA; 5914 BP.
 AC
 XX ABL34168;
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 2141.
 XX
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antihaemic; cytosinetic; neotropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW

KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 XX 02-JUL-2001; 2001WO-EP07537.
 PF
 XX 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 PI WPI: 2002-130909/17.
 XX
 DR Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 2141; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 5914 BP; 1235 A; 162 C; 1616 G; 2900 T; 1 other;
 XX
 Query Match 7.7%; Score 36.8; DB 24; Length 5914;
 Best Local Similarity 51.2%; Pred. No. 2.4;
 Matches 86; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
 XX
 QY 31 ACATGATCCTTTATCCACGCTTACAGTTTAATTAACCTGCCAGCATTAAGGATGTTT 90
 DB 3073 AACTGCCATTCACACAGCTCAATCAATTTAAACCAATACCTTTATCCATCATTC 3014
 QY 91 TACCTTAAGTCAAAAAATAGCGAATTTTCCAAAGCAAAAGCTAATATCGCAAAAC 150
 DB 3013 TACCTCTATTAATTAATTTCTTAACAACCCACGAAATACATATACCGCTCGAAAAAT 2954
 QY 151 CTCAGTAAATATCTGCTGAGCTATTTATGCTAAGTAACTTTATCC 198
 DB 2953 TCCAAACCCAAAAAATACAAAGCTATTTACATTAATAATACGTATCCC 2906
 XX
 RESULT 14
 ID ABL70321/c
 XX ABL70321 standard; DNA; 6310 BP.
 AC
 XX ABL70321;
 DT 01-JUL-2002 (first entry)
 XX
 DE Chemically treated cell signalling DNA sequence#106.
 XX
 KW Cell signalling; cytosine methylation; cell signalling disease;
 KW cancer; tumour; cytosinetic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200202807-A2.

XX 10-JAN-2002.
PD
XX
XX 29-JUN-2001; 2001WO-EP07471.
PF
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
DR WPI; 2002-154758/20.
XX
XX Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signalling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signalling -
XX
XX
PS Claim 1; SEQ ID NO 211; 24pp+sequence listing; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records AB170111-AB170626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
XX
SQ Sequence 6310 BP; 1426 A; 213 C; 1583 G; 3088 T; 0 other;
XX
XX
XX Query Match 7.7%; Score 36.8; DB 24; Length 6310;
XX Best local Similarity 54.4%; Pred. No. 2.5;
XX Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
XX
XX 92 AACTTAAGTCAAAAAATGCGCATTTTCCACGACAAAGCTAATATCGCAAAACC 151
XX 2449 AATTAACCAAAAAAATGCGCAAAATTTCCAAACAAAAACAAAAATACAAAAACC 2390
XX
XX 152 TCAGTAAATCTTGCTGGAGTATATTCCTAAGTAACATTACCCCGTGAAGTAAG 211
XX 2389 TAAAAAATATCAACTTAATATATTAATAAAAAATAACGAAAAAACCCTCAATATTAAATA 2330
XX
XX 212 GATCAATCAAGAGAGA 227
XX 2329 AACAAAAAATAAAAAA 2314
XX
XX
XX RESULT 15
XX AAS61268/C
XX ID AAS61268 standard; DNA; 6310 BP.
XX
XX AAS61268;
XX
XX 29-JAN-2002 (first entry)
XX
XX Human gene regulation-associated gene oligonucleotide #223.
XX
XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
XX cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
XX asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
XX renal disease; Preeclampsia; cardiac allograft vascular disease;
XX colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
XX immunostimulant; cardiact; antiinflammatory; coagulant; antisthmatic;
XX nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
XX

XX Homo sapiens.
OS
XX WO200177375-A2.
PN
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-EP03968.
XX
XX 06-APR-2000; 2000DE-1019058.
PR
XX 07-APR-2000; 2000DE-1019173.
PR
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
DR WPI; 2002-017470/02.
XX
XX
XX New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease -
XX
XX
XX Disclosure; SEQ ID No 229; 26pp; English.
XX
XX
XX The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC preeclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 6310 BP; 1426 A; 213 C; 1583 G; 3088 T; 0 other;
XX
XX
XX Query Match 7.7%; Score 36.8; DB 24; Length 6310;
XX Best local Similarity 54.4%; Pred. No. 2.5;
XX Matches 74; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
XX

Search completed: February 7, 2003, 19:59:38
Job time : 1034 secs

THIS PAGE BLANK (USPTO)

APPLICANT: Lavallee, Edward R.
APPLICANT: Racine, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,022
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4078 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-960-022-3

Query Match
Best Local Similarity 7.0%; Score 33.6; DB 2; Length 4078;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 53 AACAGTTTAATACCTGCCAGCAATAGGAGTGTGTTAACTTAGTCAAAAAAATAGC 112
DB 3720 AAAAGAGTCTTATCTTCTCCCAAAATAGAAATTTGTATTCCTATAGTCAAAACAATAA 3661
QY 113 GAATTTCCACGACGACAAAGCTAATATCGCAAAA 148
DB 3660 CATTTTGAATTAATAAAAGTGAGACAACTTGAAA 3625

RESULT 3
US-08-694-869-2
Sequence 2, Application US/08694869
Patent No. 5994123
GENERAL INFORMATION:
APPLICANT: Olszewski, N.
APPLICANT: Tzafir, I.
APPLICANT: Somers, D. A.
APPLICANT: Lockhart, B.
APPLICANT: Torbert, K.
TITLE OF INVENTION: SUGARCANE BACILLIFORM VIRUS
TITLE OF INVENTION: PROMOTER
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Riuth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,869
FILING DATE: 09-AUG-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 600,369US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 7568 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-694-869-2

Query Match
Best Local Similarity 6.7%; Score 32.4; DB 2; Length 7568;
Matches 63; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 53 AACAGTTTAATACCTGCCAGCAATAGGAGTGTGTTAACTTAGTCAAAAAAATAGC 112
DB 5962 AACATATTTATTAAGAGATTGCGAATTTGATGACGAGAGTTGAAGCAAAAGAGTC 6021
QY 113 GAATTTCCACGACGACAAAGCTAATATCGCAAAAACCTCAGTAAATCTTG 166
DB 6022 TTAATCTGCTGCTAGCAGACACTGACTATGCCGAAACACATCAAAAGATATGG 6075

RESULT 4
US-09-349-546-2
Sequence 2, Application US/09349546
Patent No. 6093569
GENERAL INFORMATION:
APPLICANT: Olszewski, N.
APPLICANT: Tzafir, I.
APPLICANT: Somers, D.A.
APPLICANT: Lockhart, B.
APPLICANT: Torbert, K.
TITLE OF INVENTION: Sugarcane bacilliform virus promoter
FILE REFERENCE: 600,369US2
CURRENT APPLICATION NUMBER: US/09/349,546
CURRENT FILING DATE: 1999-07-08
EARLIER APPLICATION NUMBER: US 08/694,869
EARLIER FILING DATE: 1996-08-09
EARLIER APPLICATION NUMBER: PCT/IB97/01338
EARLIER FILING DATE: 1997-08-13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 7568
TYPE: DNA
ORGANISM: sugarcane bacilliform virus
US-09-349-546-2

Query Match
Best Local Similarity 6.7%; Score 32.4; DB 3; Length 7568;
Matches 63; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 53 AACAGTTTAATACCTGCCAGCAATAGGAGTGTGTTAACTTAGTCAAAAAAATAGC 112

Db 5962 ACATATTTATTAAGAGATTGCGAATTTGATACGAGAAAGTGGAAGACAAAAGAGTC 6021
QY 113 GAATTTTCCACGACACAAAAGCTAAATATTCGCAAAAACCTCAGTAAATCTTG 166
Db 6022 TTAATTCCTGGTCAGCAACACTGACTATGCGAAGAAACATCAAGATATG 6075

RESULT 5

US-08-299-953-1
; Sequence 1, Application US/08299953
; Patent No. 5646333
; GENERAL INFORMATION:
; APPLICANT: Dobres, Michael S. and Mandact, Sevnur
; TITLE OF INVENTION: A Plant Promoter Useful for Directing the
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5646333rls
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,953
; FILING DATE: Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: NOVA-0003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-564-8960
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2861 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-299-953-1

Query Match 6.7%; Score 32; DB 1; Length 2861;

Best Local Similarity 50.7%; Pred. No. 2.4;
Matches 77; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 89 TTTAACTTAAGTCAAAAAATAGCAATTTTCCACGACACAAAAGCTAAATATCGCAAAA 148
Db 1089 TATATCTAATATTAATTAATCAATTTTCAATTAAGAAATTAATATGTTTTTATAT 1148
QY 149 ACCTCAGTAAAAATCTTGCTGAGACTATTTGCTAGTAACATTTACCCCTGAAGTTA 208
Db 1149 AACCTGGAATTTTATATATCTTTTTTTAAATTAATTAATTTATCTTCATTTT 1208
QY 209 ATGATCAATCAAGAGAGATGGGCTGTAA 240
Db 1209 ATTAATATATAAGATATTTTTTGATATAAT 1240

RESULT 6

US-08-459-415-1
; Sequence 1, Application US/08459415
; Patent No. 5744334
; GENERAL INFORMATION:
; APPLICANT: Dobres, Michael S. and Mandact, Sevnur

; TITLE OF INVENTION: A Plant Promoter Useful for Directing the
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5744334rls
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; ZIP: 19103

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,415
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/299,953

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: NOVA-0003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-564-8960
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2861 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-459-415-1

Query Match 6.7%; Score 32; DB 1; Length 2861;

Best Local Similarity 50.7%; Pred. No. 2.4;
Matches 77; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 89 TTTAACTTAAGTCAAAAAATAGCAATTTTCCACGACACAAAAGCTAAATATCGCAAAA 148
Db 1089 TATATCTAATATTAATTAATCAATTTTCAATTAAGAAATTAATATGTTTTTATAT 1148
QY 149 ACCTCAGTAAAAATCTTGCTGAGACTATTTGCTAGTAACATTTACCCCTGAAGTTA 208
Db 1149 AACCTGGAATTTTATATATCTTTTTTTAAATTAATTAATTTATCTTCATTTT 1208
QY 209 ATGATCAATCAAGAGAGATGGGCTGTAA 240
Db 1209 ATTAATATATAAGATATTTTTTGATATAAT 1240

RESULT 7

US-09-066-687-1
; Sequence 1, Application US/09066687
; Patent No. 6339185
; GENERAL INFORMATION:
; APPLICANT: Dobres, Michael S. and Mandact, Sevnur
; TITLE OF INVENTION: A Plant Promoter Useful for Directing the
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6339185rls
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
; STATE: PA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11231
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/299,953
FILING DATE: September 2, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: NOVA-0016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-564-8960
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3881 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-11231-2

Query Match 6.7%; Score 32; DB 5; Length 3881;
Best Local Similarity 50.7%; Pred. No. 2.8;
Matches 77; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 89 TTTACTTAACTCAAAAATAGCAAGATTTTCCAAAGCAAAAGCTTAATATCGCAAAA 148
DB 1089 TATATCATATATATATATATATTTTACAAATGTTAAAGAAATTAATATTTATTTATTTAT 1148
QY 149 ACCCTAGTAAATCTTGCTGAGCTATTTATGCTAAGTAACTTACCCCTGAAGTTA 208
DB 1149 AACTCTAAGATTTTATTTATTTATTTTAAATATATATTAATTAATTTCTCCATTTT 1208
QY 209 ATGATCAATCAAGAGAGATGTGGCGCTGTAA 240
DB 1209 ATTAATATATTAAGATATTTTGTGATATTAAT 1240

RESULT 13
US-08-341-568-1
Sequence 1, Application US/08341568
Patent No. 5661021
GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Viikari, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Saloheimo, Anu
APPLICANT: Marjatta, Rannu
TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching lig
TITLE OF INVENTION: pulps
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US

ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/341,568
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr, Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 365-262P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 536 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Trichoderma reesei
STRAIN: QM9414
US-08-341-568-1

Query Match 6.6%; Score 31.8; DB 1; Length 536;
Best Local Similarity 50.3%; Pred. No. 1.3;
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 18 TGATGCTATTTGACACTATTCCTTACCAGCTCAACAGTTTAATTAACCTGCGACCAAT 77
DB 382 TGATCTCTGTAACATTCATTCATTAAGTAAACCTTCTTAATGATTAAGCTAGAAC 441
QY 78 AAGGATGTTGTTTAACTTAAGTCAAAAAAATAGCAATTTTCCAAAGCAAAAGCTAA 137
DB 442 AAAGTTTGAGGAGTAAAGTAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAA 501
QY 138 ATATCGCAAAAGCTCACTAAATCTTGCTGAG 172
DB 502 AAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAAAAGCAAAAA 536

RESULT 14
US-08-911-020-1
Sequence 1, Application US/08911020
Patent No. 5854047
GENERAL INFORMATION:
APPLICANT: Buchert, Johanna
APPLICANT: Silka-aho, Matti
APPLICANT: Viikari, Liisa
APPLICANT: Penttila, Merja
APPLICANT: Saloheimo, Anu
APPLICANT: Marjatta, Rannu
TITLE OF INVENTION: Mannanase enzymes, genes coding for them,
TITLE OF INVENTION: methods for isolating the genes, and methods for bleaching
TITLE OF INVENTION: pulps
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: PO Box 747
CITY: Falls Church
STATE: VA
COUNTRY: US
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

1      SOFTWARE: PatentIn Release #1.0, Version #1.30
2
3      CURRENT APPLICATION DATA:
4
5      APPLICATION NUMBER: US/08/911,020
6
7      FILING DATE: 13-AUG-1997
8
9      CLASSIFICATION: 435
10
11     PRIOR APPLICATION DATA:
12
13     APPLICATION NUMBER: US 08/341,568
14
15     FILING DATE: 22-NOV-1994
16
17     ATTORNEY/AGENT INFORMATION:
18
19     NAME: Murphy Jr, Gerald M
20
21     REGISTRATION NUMBER: 28,977
22
23     REFERENCE/DOCKET NUMBER: 365-262P
24
25     TELECOMMUNICATION INFORMATION:
26
27     TELEPHONE: (703) 205-8000
28
29     TELEFAX: (703) 205-8050
30
31     TELEEX: 248345
32
33     INFORMATION FOR SEQ ID NO: 1:
34
35     SEQUENCE CHARACTERISTICS:
36
37     LENGTH: 536 base pairs
38
39     TYPE: nucleic acid
40
41     STRANDEDNESS: single
42
43     TOPOLOGY: linear
44
45     MOLECULE TYPE: cDNA
46
47     ORIGINAL SOURCE:
48
49     ORGANISM: Trichoderma reesei
50
51     STRAIN: QM9414
52
53     IS-08-911-020-1

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 19:42:18 ; Search time 98 Seconds

(without alignments)
2326.484 Million cell updates/sec

Title: US-09-828-456-1

Perfect score: 481

Sequence: 1 tgcctctcattgaagatga.....gtgatgtcagagatagtaa 481

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 481 | 100.0 | 481 | 10 | US-09-828-456-1 |
| 2 | 38.6 | 8.0 | 1947 | 9 | US-09-938-842A-3476 |
| 3 | 36.8 | 7.7 | 332 | 10 | US-09-983-965-5260 |
| 4 | 36.8 | 7.7 | 7053 | 10 | US-09-764-864-7292 |
| 5 | 35.6 | 7.4 | 443 | 10 | US-09-876-889-272 |
| 6 | 35 | 7.3 | 66109 | 10 | US-09-880-107-3768 |
| 7 | 35 | 7.3 | 198285 | 10 | US-09-880-107-3768 |
| 8 | 34.2 | 7.1 | 1512 | 9 | US-09-938-842A-3316 |
| 9 | 34.2 | 7.1 | 2000 | 10 | US-09-887-576-859 |
| 10 | 34.2 | 7.1 | 2075 | 9 | US-10-098-841-110 |
| 11 | 33.6 | 7.0 | 4083 | 10 | US-09-925-302-177 |
| 12 | 33.4 | 6.9 | 241 | 10 | US-09-960-352-7904 |
| 13 | 33 | 6.9 | 418 | 10 | US-09-867-701-2763 |
| 14 | 33 | 6.9 | 335913 | 9 | US-09-754-853A-2 |
| 15 | 33 | 6.9 | 335913 | 9 | US-09-754-853A-3 |
| 16 | 32.6 | 6.8 | 683 | 10 | US-09-764-877-914 |
| 17 | 32.6 | 6.8 | 2000 | 9 | US-09-938-842A-5219 |
| 18 | 32.4 | 6.7 | 2846 | 9 | US-10-008-016-1 |
| 19 | 32.4 | 6.7 | 7200 | 9 | US-09-853-450-48 |

| | | | | | | | |
|---|----|------|-----|-------|----|---------------------|---------------------|
| C | 20 | 32.2 | 6.7 | 1142 | 9 | US-10-098-841-160 | Sequence 160, App |
| C | 21 | 32.2 | 6.7 | 19553 | 10 | US-09-764-847-1425 | Sequence 1425, App |
| C | 22 | 31.8 | 6.6 | 411 | 10 | US-09-867-701-7321 | Sequence 7321, App |
| C | 23 | 31.8 | 6.6 | 516 | 10 | US-09-960-352-5785 | Sequence 5785, App |
| C | 24 | 31.6 | 6.6 | 372 | 10 | US-09-783-590-1292 | Sequence 1292, App |
| C | 25 | 31.6 | 6.6 | 1902 | 10 | US-09-349-385-2 | Sequence 2, App1 |
| C | 26 | 31.6 | 6.6 | 2418 | 10 | US-09-925-300-719 | Sequence 719, App |
| C | 27 | 31.6 | 6.6 | 7869 | 10 | US-09-349-385-3 | Sequence 3, App1 |
| C | 28 | 31.6 | 6.6 | 12886 | 9 | US-10-114-170-14 | Sequence 14, App1 |
| C | 29 | 31.4 | 6.5 | 4239 | 10 | US-09-815-048-1 | Sequence 1, App1 |
| C | 30 | 31.2 | 6.5 | 275 | 10 | US-09-867-701-10310 | Sequence 10310, App |
| C | 31 | 31 | 6.4 | 2000 | 9 | US-09-938-842A-2799 | Sequence 2799, App |
| C | 32 | 31 | 6.4 | 70768 | 9 | US-10-135-322-13 | Sequence 13, App1 |
| C | 33 | 30.8 | 6.4 | 368 | 10 | US-09-834-975-144 | Sequence 144, App |
| C | 34 | 30.8 | 6.4 | 5444 | 10 | US-09-896-617-1 | Sequence 1, App1 |
| C | 35 | 30.8 | 6.4 | 5444 | 10 | US-09-931-071-1 | Sequence 1, App1 |
| C | 36 | 30.8 | 6.4 | 32193 | 9 | US-09-764-868-1508 | Sequence 1508, App |
| C | 37 | 30.8 | 6.4 | 98865 | 10 | US-09-770-689A-3 | Sequence 3, App1 |
| C | 38 | 30.6 | 6.4 | 392 | 10 | US-09-833-381-1132 | Sequence 1132, App |
| C | 39 | 30.6 | 6.4 | 415 | 10 | US-09-867-701-2926 | Sequence 2926, App |
| C | 40 | 30.6 | 6.4 | 520 | 10 | US-09-998-598-434 | Sequence 434, App |
| C | 41 | 30.6 | 6.4 | 814 | 12 | US-10-001-870-42 | Sequence 42, App1 |
| C | 42 | 30.6 | 6.4 | 1205 | 12 | US-10-001-870-43 | Sequence 43, App1 |
| C | 43 | 30.6 | 6.4 | 2275 | 10 | US-09-764-864-209 | Sequence 209, App1 |
| C | 44 | 30.6 | 6.4 | 5232 | 10 | US-09-974-592-3 | Sequence 3, App1 |
| C | 45 | 30.6 | 6.4 | 8726 | 10 | US-09-070-927A-95 | Sequence 95, App1 |

ALIGNMENTS

| | | | |
|-----------------------------|---|---|-------------|
| RESULT 1 | US-09-828-456-1 | Sequence 1, Application US/09828456 | |
| Patent No. | US20020051982A1 | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: | Levy, Stewart et al. | | |
| TITLE OF INVENTION: | No. US20020051982A1 BLR Molecules affecting Antibiotic Susc | | |
| FILE REFERENCE: | PRZ-030 | | |
| CURRENT APPLICATION NUMBER: | US/09/828, 456 | | |
| CURRENT FILING DATE: | 2001-04-06 | | |
| PRIOR APPLICATION NUMBER: | 60/195,505 | | |
| PRIOR FILING DATE: | 2000-04-06 | | |
| PRIOR APPLICATION NUMBER: | 60/218,380 | | |
| PRIOR FILING DATE: | 2000-07-04 | | |
| NUMBER OF SEQ ID NOS: | 2 | | |
| SOFTWARE: | Patentlin Ver. 2.0 | | |
| SEQ ID NO 1 | | | |
| LENGTH: | 481 | | |
| TYPE: | DNA | | |
| ORGANISM: | Escherichia coli | | |
| FEATURE: | | | |
| NAME/KEY: | CDS | | |
| LOCATION: | (239)..(361) | | |
| Query Match | 100.0%; | Score 481; | DB 10; |
| Best Local Similarity | 100.0%; | Pred. No. 5.3e-131; | Length 481; |
| Matches 481; | Conservative 0; | Mismatches 0; | Indels 0; |
| Gaps 0; | | | |
| QY | 1 | TGCCTCATGAGTATGATGCTATTTGACACTATCTTACCCACGCTCAACAGTTT | 60 |
| DB | 1 | TGCCTCATGAGTATGATGCTATTTGACACTATCTTACCCACGCTCAACAGTTT | 60 |
| QY | 61 | AATAACCTGCCAGCAATTAAGGATGTGTTAACTTAAGTCAAAAAATACGATTTTC | 120 |
| DB | 61 | AATAACCTGCCAGCAATTAAGGATGTGTTAACTTAAGTCAAAAAATACGATTTTC | 120 |
| QY | 121 | CAACGACAAAGCTAATATCGCAAAACCTCGTAATAATCTTCTGCGACTATATT | 180 |
| DB | 121 | CAACGACAAAGCTAATATCGCAAAACCTCGTAATAATCTTCTGCGACTATATT | 180 |
| QY | 181 | GCTAAGTAACTTATACCCCTGAAGTTAATGATCAATCAAGAGATGTGGCTGTAT | 240 |

```
|||||
Db 181 GCTAAGTAATTTACCCCTGAAGTAAATGATCAATCAAGAGATGCGCTGAAT 240
QY 241 GAATGCTCTTATGAATTAACAGGTGATGCTTCTTCTTCTTCTTCTTCTTCTG 300
Db 241 GAATGCTCTTATGAATTAACAGGTGATGCTTCTTCTTCTTCTTCTTCTTCTG 300
QY 301 CGTGGCAGTCACTTATGACATATGACGACCTGGAACAGAGCTTCTTCTTCTTCTG 360
Db 301 CGTGGCAGTCACTTATGACATATGACGACCTGGAACAGAGCTTCTTCTTCTTCTG 360
QY 361 GTAGCTCTGACTTGTGAGCAGCATGCTCCGCTCCGGGTGATGCTTCTTCTTCTG 420
Db 361 GTAGCTCTGACTTGTGAGCAGCATGCTCCGCTCCGGGTGATGCTTCTTCTTCTG 420
QY 421 GATATTAACCGAGATTTATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 480
Db 421 GATATTAACCGAGATTTATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 480
QY 481 A 481
Db 481 A 481
```

RESULT 2

```
US-09-938-842A-3476/C
: Sequence 3476, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kieps, Joel
: APPLICANT: Mang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPI300-3
: CURRENT APPLICATION NUMBER: US/09/938, 842A
: PRIOR FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 3476
: LENGTH: 1947
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-3476
```

```
Query Match 8.0%; Score 38.6; DB 9; Length 1947;
Best Local Similarity 48.8%; Pred. No. 0.13;
Matches 104; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
```

```
QY 16 TATGATGCTATTGACACTATCTTACCAAGCTCAACAGTTAATTAACCTGCCAGCA 75
Db 966 TATGATGCTATTGACACTATCTTACCAAGCTCAACAGTTAATTAACCTGCCAGCA 75
QY 76 ATAAGGATGCTTGTACTTACTTAAGCAAAAATATGCAATTTTCCAGCAAGAAAGCT 135
Db 906 TTAAGGATGCTTGTACTTACTTACTTAAGCAAAAATATGCAATTTTCCAGCAAGAAAGCT 135
QY 136 AATATGCGCAAAACCTGAGTAAATATCTTGTGAGCTATATATGCTAAGTAACTTTA 195
Db 846 TAAATATGCTTGTACTTACTTACTTAAATATATGATTTGAGTAAATATGCTAAGTAACTTTA 195
QY 196 CCCCTGAAGTTAATGATCAATCAAGAGAT 228
Db 786 ACCCATTTTCAATCATGGAACAAATTAAGTT 754
```

RESULT 3

```
US-09-983-965-5260/C
: Sequence 5260, Application US/09983965
: Patent No. US20020137160A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengding
: APPLICANT: Byatt, John C.
: APPLICANT: Mathalagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
: FILE REFERENCE: 37-21(10297)C
: CURRENT APPLICATION NUMBER: US/09/983,965
: CURRENT FILING DATE: 2001-10-26
: PRIOR APPLICATION NUMBER: US 09/465,231
: PRIOR FILING DATE: 1999-12-15
: PRIOR APPLICATION NUMBER: US 60/113,678
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 5912
: SEQ ID NO 5260
: LENGTH: 332
: TYPE: DNA
: ORGANISM: Bos taurus
: FEATURE:
: OTHER INFORMATION: Clone ID: 38-LIB34-041-Q1-E1-B6
US-09-983-965-5260
```

```
Query Match 7.7%; Score 36.8; DB 10; Length 332;
Best Local Similarity 51.9%; Pred. No. 0.2;
Matches 83; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```
QY 103 AAAAAATAGCGAATTTTCCACACACAAAGCTAAATATGCGCAAAACCTGATTAAT 162
Db 212 AATTAATTTATCACTTATCTCAGAAAATATTAATTAACAAAGTAAATTTAAAGCCAAATA 153
QY 163 CTGCTGAGCTATTTATGCTAATGATTAATTTACCCCTGAAGTTAATGATCAATCAAG 222
Db 152 TTTCACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 93
QY 223 AGAGATGTGGCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 262
Db 92 ACATTAGAGATTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53
```

RESULT 4

```
US-09-764-864-792/C
: Sequence 792, Application US/09764864
: Patent No. US20020132753A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT23
: CURRENT APPLICATION NUMBER: US/09/764,864
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1792
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 792
: LENGTH: 7053
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (154)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-792
```

```
Query Match 7.7%; Score 36.8; DB 10; Length 7053;
Best Local Similarity 51.2%; Pred. No. 0.79;
Matches 86; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
```

```
QY 90 TTAAGTCAAAAAATAGCAATTTTCCAAAGCAAAAGCTAAATATGCGCAAAA 149
Db 3528 TTAAGTCAAAAAATAGCAATTTTCCAAAGCAAAAGCTAAATATGCGCAAAA 3469
```



```
QY 150 CCTCAGTAAATCTGCTGAGCTATTATGCTAAGTACATTACCCCTGAGTTAA 209
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3468 ACAAAACAAACAAACAGAGAGAGATCATTAACCACTGATATATTTGTATATAA 3409
QY 210 TGGATCATCAAGAGAGATGTGGCTGTAATGAATCCTCTTAATGAAT 257
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3408 TAGCTCAGTGTAAAGAAATATGATTTGAGTGGTGCATGCAAGTAAC 3361
```

RESULT 5

```
US-09-876-889-272
; Sequence 272, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C3
; CURRENT APPLICATION NUMBER: US/09/876,889
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 272
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(443)
; OTHER INFORMATION: n = A,T,C or G
US-09-876-889-272
```

```
Query Match
Best Local Similarity 7.4%; Score 35.6; DB 10; Length 443;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 100 TCAAAAAATAGCGATTTTCCACAGACAAAGCTAAATATGCGCAAAACCTCAGTAA 159
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 207 TCAAAACAAATACAAATATCTCTGAGAAAGAAATTTAAGAACTTAATTAAGTAAA 266
QY 160 AATCTCTGAGCTATTTATGCTAATACATTACCCCTGAGTTATGATCATC 219
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 267 AATGCTTTCTGAAATATCTTGGAAATTTTAAATGTCAAAATGATGACTCATGC 326
QY 220 AAGAGAGATGTGGCTCT 237
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 327 TAAATACATTGAGGGCT 344
```

RESULT 6

```
US-09-880-107-3768
; Sequence 3768, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3768
```

```
; LENGTH: 66109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X66401
; NAME/KEY: unsure
; LOCATION: (1)..(66109)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3768
```

```
Query Match
Best Local Similarity 7.3%; Score 35; DB 10; Length 66109;
Matches 101; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
```

```
QY 13 AAGTATGATGGCTATTGTGACACTATCTTACCACGCTCAACAGTTAATTAACCTGCA 72
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13691 AAGATATTTCTTAATCTCGCCTCTCTTAATCCATCCATGCAATGAACTCTTTA 13750
QY 73 GCATTAAGGATGTGTTTAACTTAAGTCAAAAAATAGGAAATTTCCACGACAA 132
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13751 ATTATATTTCTGTTATTTTATTTGAATGGAATAATATATATCAATTTGTCA 13810
QY 133 GCTAAATATGCAAAAACCTCAGTAAATAATCTGCTGAGCTATTATGCTAAGTAACAT 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13811 TGGAAATATAGGAATATCTCTCAGAGATATGATGTTATTAAGTCATTGACATATTG 13870
QY 193 TTACCCCTGAGTTAATGATCANTCAGA 223
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13871 TTAACTCAGATGCAAAATGCTGGAATTAATA 13901
```

RESULT 7

```
US-09-880-107-3814
; Sequence 3814, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3814
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X87344
US-09-880-107-3814
```

```
Query Match
Best Local Similarity 7.3%; Score 35; DB 10; Length 198285;
Matches 101; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
```

```
QY 13 AAGTATGATGGCTATTGTGACACTATCTTACCACGCTCAACAGTTAATTAACCTGCA 72
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 89480 AAGATATTTCTTAATCTCGCCTCTCTTAATCCATCCATGCAATGAACTCTTTA 89539
QY 73 GCATTAAGGATGTGTTTAACTTAAGTCAAAAAATAGGAAATTTCCACGACAA 132
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 89540 ATTATATTTCTGTTATTTTATTTGAATGGAATAATATATATCAATTTGTCA 89599
QY 133 GCTAAATATGCAAAAACCTCAGTAAATAATCTGCTGAGCTATTATGCTAAGTAACAT 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 89600 TGGAAATATAGGAATATCTCTCAGAGATATGATGTTATTTAAGTCATTGACATATTG 89659
```

| | | | |
|----|-------|--------------------------------|-------|
| Oy | 193 | TTACCCCTGAAGTTAATGCATCAACAGA | 223 |
| | | | |
| Db | 89660 | TTAACCTCAGATGCAAAATGGTGAATTATA | 89690 |

RESULT 8

US-09-938-842A-3316/C
: Sequence 3316, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPI300-3
: CURRENT APPLICATION NUMBER: US/09/938, 842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227, 866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264, 647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300, 111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 3316
: LENGTH: 1312
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-3316

| | | | | |
|--------------------------|-------|---------------|------|-------------------|
| Query Match | 7.1% | Score 34.2 | DB 9 | Length 1512 |
| Best Local Similarity | 51.7% | Pred. No. 2.2 | | |
| Matches 78; Conservative | 0 | Mismatches | 73 | Indels 0; Gaps 0; |

| QY | 93 | ACTTACGTCAAAAAATATGCGCAATTTTTCACAGCAACAAAGCTAATATATGCGAAAAA | 152 |
|----|-----|---|-----|
| Db | 510 | AATTTATTCAGAGTTATCTACAAATTTTTCACCTAAGAAAACCTATCTATGTGACATTC | 451 |
| QY | 153 | CAGTAAAAATCTTGCTGGAGCTATTATTGCTAAGTAACATTTTACCCTGAGATTATGG | 212 |
| Db | 450 | TTGAACCAATCTATATACATACCTAGTTGTCTATTTGTAGTTTCGCGAAAGCAATGGAA | 391 |
| QY | 213 | ATTCATCAAGAGAGATGTGGCTGTAATGAA | 243 |
| Db | 390 | AACATCTAGTGAGATTTGTCTCTTACAA | 360 |

RESULT 5

```

US-09-887-576-859
: Sequence 859. Application US/09887576
: Patent No. US20020144047A1
:
: GENERAL INFORMATION:
: APPLICANT: Budworth, P.
: APPLICANT: Brown, D.
: APPLICANT: Chang, H.
: APPLICANT: Zhu, T.
: APPLICANT: Han, B.
: APPLICANT: Wang, X.
: APPLICANT: Cooper, Bret
: TITLE OF INVENTION: Promoters for regulation of plant expression
: FILE REFERENCE: 1360.001US1
: CURRENT APPLICATION NUMBER: US/09/887, 576
: CURRENT FILING DATE: 2001-06-25
: PRIOR APPLICATION NUMBER: US 60/213, 848
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: US 60/214, 087
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: US 60/258, 692
: PRIOR FILING DATE: 2000-12-29
: NUMBER OF SEQ ID NOS: 875
:

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 859
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-859

```

| | | | | |
|-----------------------|--------------|--------------|---------------|-------------|
| Query Match | 7.1% | Score 34.2 | DB 10 | Length 2000 |
| Best Local Similarity | 56.8% | Pred No. 2.6 | | |
| Matches 63 | Conservative | 0 | Mismatches 48 | Indels 0 |
| | | | | Gaps 0 |

| Accession | Sequence | Position |
|-----------|--|----------|
| Q7 | TTTTTAACTAAGCAAAAAATAGCGATTTTCCACGCAAAAGCTAAATATGCA | 1 45 |
| Db | TTGACTAATTTTTCGAAAAAATAGTAAATTTTAAACGAGCAAAATATATTGAAA | 5 48 |
| Q7 | AAACCTCAGTAAAAATCTGCTGAGCTATTATGCTAGTAACATTTC | 196 |
| Db | AATTATTAATATTATGATTATATAAACTATGCTGGTGTGTATAATATTC | 599 |

RESULT 10

```

US-10-098-841-110
? Sequence 110, Application US/10098841
? Publication No. US20020197679A1
?
? GENERAL INFORMATION:
? APPLICANT: Tang, Y. Tom
? APPLICANT: Liu, Chenghua
? APPLICANT: Asundi, Vinod
? APPLICANT: Xu, Chongjun
? APPLICANT: Zhou, Ping
? APPLICANT: Ma, Yundong
? APPLICANT: Wang, Jian-Rui
? APPLICANT: Zhao, Qing A.
? APPLICANT: Ren, Feiyan
? APPLICANT: Chen, Rui-hong
? APPLICANT: Wang, Dunrui
? APPLICANT: Wang, Zhiwei
? APPLICANT: Weinman, Tom
? APPLICANT: Zhang, Jie
? APPLICANT: Qian, Xiaohong B.
? APPLICANT: Drmanac, Radote T.
? TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
? TITLE OF INVENTION: Polypeptides
?
? FILE REFERENCE: 784C1f2
? CURRENT APPLICATION NUMBER: US/10/098, 841
? CURRENT FILING DATE: 2002-03-13
? PRIOR APPLICATION NUMBER: 09/598, 042
? PRIOR FILING DATE: 2000-06-20
? PRIOR APPLICATION NUMBER: 09/552, 317
? PRIOR FILING DATE: 2000-04-25
? PRIOR APPLICATION NUMBER: 09/488, 725
? PRIOR FILING DATE: 2000-01-21
? NUMBER OF SEQ ID NOS: 331
? SOFTWARE: pl_FL_genes Version 1.0
? SEQ ID NO 110
?
? LENGTH: 2075
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (147)..(2075)
US-10-098-841-110

```

| | | | | |
|-----------------------|--------|----------------|-------|----------------|
| Query Match | 7.18; | Score 34.2; | DB 9; | length 2075; |
| Best Local Similarity | 50.98; | Pred. No. 2.6; | | |
| Matches | 81; | Conservative | 0; | Mismatches 78; |
| | | | | Indels 0; |
| | | | | Gaps 0; |

QY 56 AGTTTAATACTGCGACGAATAAGGAGTGTGGTTTAACTTAAGCAAAAAAATAGCGAA 115
| | | | | | | | | | | | | | | |
Db 499 ATTTTGAAATGACTCAGATATTATACAATTTGAGGCGATCAGGAGATCCAAAGAAGAA 558
| | | | | | | | | | | | | | | |
QY 116 TTTTTCAGGACAAAGCTAATATGCGAAAACCCTAGTAAAAATCTTGCTGCAGCTA 175
| | | | | | | | | | | | | | | |

Db 559 ATATCAATCAAAAGATGATCATGCTATGAAAAAGCTGCTACTCATCTCCTCTCTC 618
Qy 176 TTATTGCTAGTACATTTTACCCTCGAAGTTAATGAT 214
Matches 61: Conservative 0; Mismatches 46; Indels 0; Gaps 0;
Db 619 TTATTGCAATACATTAACCCGTATGATGTAAGAAAT 657

RESULT 11
US-09-925-302-177/c

; Sequence 177, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 177
; LENGTH: 4083
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (48)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (157)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-177

Query Match 7.0%; Score 33.6; DB 10; Length 4083;
Best Local Similarity 59.4%; Pred. No. 5.3;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 53 AACGTTTAATACCTGCCAGCATTAAGGATGTTGTTAATCAAGCAAAAAATAGC 112
Db 3725 AAAAGAGTCTTATCTCCAAATAGCAAAATTTGATTCCTATAGCAAAATTA 3666
Qy 113 GAATTTCCACGAACAAAGCTAAATTCGCAAA 148
Db 3665 CAATTTTGAATTAATAAAAGTGAAGATCTTGAAA 3630

RESULT 12
US-09-960-352-7904/c
; Sequence 7904, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7904
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 34-LIB3057-019-Q1-K1-A10
US-09-960-352-7904

Query Match 6.9%; Score 33.4; DB 10; Length 241;
Best Local Similarity 57.0%; Pred. No. 1.7;
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 55 CAGTTTAATACCTGCCAGCAATTAAGGATGTGTTTAACCTAAGTCAAAAAATAGCA 114
Db 224 CATATTAATTAATTAAGTAAGTAATGATTTTTTTTAAAAAAATTTAAAAAATTAATA 165
Qy 115 ATTTTCCACGACAAAGCTAAATATCGCAAAACCTCAGTAAAA 161
Db 164 AAATTAATATATAATATAAAAAAAAAAAAAAAAAAAAAA 118

RESULT 13

US-09-867-701-2763
; Sequence 2763, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2763
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-2763

Query Match 6.9%; Score 33; DB 10; Length 418;
Best Local Similarity 58.8%; Pred. No. 2.8;
Matches 57; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 81 GGATGTGTTTAAGTCAAGCAATTAAGTCAAGCAATTAAGTCAAGCAAGCAAGTAAATA 140
Db 216 GCATTAATTTTAAGCTTTGGAATTAATTAATCTATTCCTCCGCCCAAAAAAGTA 275
Qy 141 TCGCAAAACCTCAGTAAAAATCTGCTGAGCTATT 177
Db 276 TTGTAATACCTTAAGTAAATTAATGATTTCTAGCTATT 312

RESULT 14

US-09-754-853A-2/c
; Sequence 2, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 2
; LENGTH: 335913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45163)..(45314)..(45450)..(45509)..(46941)..(48763)..(48975)..(49573)
; OTHER INFORMATION: Clone ID: 240017_region_63

Gencode version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2003, 18:32:14 ; Search time 2441 Seconds

(without alignments)
3191.327 Million cell updates/sec

Title: US-09-828-456-1

Sequence: 1 tgcctctcttgaagtatga.....gtgattgtcagtgatgtaa 481

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| c 1 | 45 | 9.4 | 672 | AA572059 | AA572059 vk95d03.r |
| 2 | 41.8 | 8.7 | 615 | AV846478 | AV846478 AV846478 |
| 3 | 41.8 | 8.7 | 712 | AV861290 | AV861290 AV861290 |
| 4 | 41.6 | 8.6 | 965 | AL105557 | AL105557 Drosophila |
| c 5 | 41.2 | 8.6 | 407 | BJ364861 | BJ364861 BJ364861 |
| 6 | 40.2 | 8.4 | 501 | AV867201 | AV867201 AV867201 |

| | | | | | | |
|------|------|-----|------|----|----------|---------------------|
| 7 | 39 | 8.1 | 940 | 17 | BH146288 | BH146288 ENTpN09TF |
| 8 | 38.2 | 7.9 | 663 | 17 | AZ347477 | AZ347477 1M0083005 |
| c 9 | 38.2 | 7.9 | 891 | 10 | BE531179 | BE531179 601278311 |
| 10 | 38 | 7.9 | 452 | 14 | H90453 | H90453 yv01c05.r1 |
| 11 | 38 | 7.9 | 633 | 13 | BJ370199 | BJ370199 BJ370199 |
| 12 | 38 | 7.9 | 864 | 17 | BH161244 | BH161244 ENTpR81TF |
| c 13 | 38 | 7.9 | 949 | 17 | AZ547495 | AZ547495 ENTpEM44TR |
| 14 | 37.8 | 7.9 | 485 | 12 | BF083668 | BF083668 M2-CT045 |
| 15 | 37.8 | 7.9 | 513 | 17 | BH288998 | BH288998 CH230-2L9 |
| c 16 | 37.6 | 7.9 | 926 | 17 | BH134032 | BH134032 ENTpNR72TF |
| 17 | 37.6 | 7.8 | 630 | 10 | BB631135 | BB631135 BB631135 |
| c 18 | 37.6 | 7.8 | 661 | 10 | AV721985 | AV721985 AV721985 |
| c 19 | 37.6 | 7.8 | 878 | 17 | AZ534410 | AZ534410 ENTpR727TF |
| c 20 | 37.4 | 7.8 | 637 | 10 | AV404587 | AV404587 AV404587 |
| c 21 | 37.4 | 7.8 | 683 | 17 | AZ660149 | AZ660149 1M0538L02 |
| c 22 | 37.4 | 7.8 | 923 | 17 | CNS0050X | AL061385 Drosophila |
| c 23 | 37.2 | 7.7 | 258 | 14 | BO601966 | BO601966 M1-P-HO-a |
| c 24 | 37.2 | 7.7 | 521 | 17 | AO181233 | AO181233 HS-3216.B |
| c 25 | 37 | 7.7 | 570 | 13 | BJ368191 | BJ368191 BJ368191 |
| 26 | 37 | 7.7 | 619 | 10 | AW180813 | AW180813 MGA0974E |
| c 27 | 37 | 7.7 | 797 | 12 | BG434608 | BG434608 602506674 |
| c 28 | 37 | 7.7 | 1000 | 17 | CNS00C00 | AL059446 Drosophila |
| c 29 | 36.8 | 7.7 | 292 | 9 | AI383833 | AI383833 tc99c03.x |
| c 30 | 36.8 | 7.7 | 295 | 10 | AW515536 | AW515536 qm85c07.x |
| c 31 | 36.8 | 7.7 | 333 | 9 | AI291645 | AI291645 qm85c07.x |
| c 32 | 36.8 | 7.7 | 341 | 9 | AI419275 | AI419275 t121h04.x |
| c 33 | 36.8 | 7.7 | 372 | 14 | H54495 | H54495 yq89f08.r1 |
| c 34 | 36.8 | 7.7 | 382 | 10 | AW576565 | AW576565 U1-HF-BNO |
| c 35 | 36.8 | 7.7 | 394 | 9 | AI361450 | AI361450 qv43e11.x |
| c 36 | 36.8 | 7.7 | 397 | 9 | AA664128 | AA664128 ac06c07.s |
| c 37 | 36.8 | 7.7 | 421 | 9 | AA983620 | AA983620 on44b02.s |
| c 38 | 36.8 | 7.7 | 426 | 9 | AI140942 | AI140942 qm83h02.s |
| c 39 | 36.8 | 7.7 | 429 | 9 | AI768981 | AI768981 wq31c03.x |
| c 40 | 36.8 | 7.7 | 439 | 9 | AI818334 | AI818334 wk59d01.x |
| c 41 | 36.8 | 7.7 | 440 | 9 | AI985789 | AI985789 wt30f04.x |
| c 42 | 36.8 | 7.7 | 440 | 9 | AA912460 | AA912460 cm52c06.s |
| c 43 | 36.8 | 7.7 | 449 | 9 | AI423972 | AI423972 t136g07.x |
| c 44 | 36.8 | 7.7 | 455 | 9 | AI802215 | AI802215 t336b06.x |
| c 45 | 36.8 | 7.7 | 466 | 9 | AA700428 | AA700428 zj74a10.s |

ALIGNMENTS

RESULT 1
AA572059/c
LOCUS
DEFINITION
vk95d03.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
36.8 KD PROTEIN IN GNTR-GGT INTERGENIC REGION ;, mRNA sequence.

ACCESSION
AA572059
VERSION
AA572059.1 GI:2345254

KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus

REFERENCE
AUTHORS
Marras M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
1 (Pages 1 to 672)
Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris B.,
Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,
Theising B., Wylie T., Lennon G., Soares B., Wilson R. and
Waterston R.

TITLE
JOURNAL
COMMENT
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL ; contact the

Db 358 AGTGAAGTCAAGAAATTAGTTCATTAATTAAGCAAGATGATTGATCTAC 417
 QY 249 TTAT 253
 Db 418 GTCTT 422

RESULT 4
 CDS015KB
 LOCUS
 DEFINITION CDS015KB 965 bp DNA linear GSS 26-JUL-1999
 Drosophila melanogaster genome survey sequence T7 end of BAC
 BACN14N18 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.

ACCESSION AL105557
 VERSION AL105557.1 GI:5617571
 KEYWORDS
 SOURCE Drosophila melanogaster.
 ORGANISM Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 965)

REFERENCE
 AUTHORS Direct Submission
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billaud at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobAC11.

FEATURES
 source
 Location/Qualifiers
 1..965
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="BACN14N18"
 /clone_lib="DrosBAC"
 /plasmid="pBelobAC11"
 /note="end : 17"
 /note="end : 17"

BASE COUNT 347 a 125 c 31 g 228 t 234 others

ORIGIN

Query Match 8.6%; Score 41.6; DB 17; Length 965;
 Best Local Similarity 44.1%; Pred. No. 1.1;
 Matches 105; Conservative 9; Mismatches 124; Indels 0; Gaps 0;

QY 61 AATPACCTGCAGCAATPAGGAGTGTCTTAACCTTAAGCAAAAAATAGCGAATTTTC 120
 Db 82 AAAAAAATNAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 141
 QY 121 CAACGAACAAGCTAAATATCGCAAAAAACCTCAGTAAATATCTCGTGAGCTATATT 180
 Db 142 WAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 201
 QY 181 GCTAAGTAACTTATCCCTCGAAGTTAATGATCAATCAAGAGATGGCGCTTAAT 240
 Db 202 TGTAAASATASASAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 261
 QY 241 GAATCGCTATGATTAACAGGTGATCGTCTTGCTGTTCACTCAATCTCTTCT 298
 Db 262 AAAAAAATTTGTTTATCTCTAATATBWKTTANTTACCATTTNCTTACTAAMT 319

RESULT 5
 LOCUS BJ364861 407 bp mRNA linear EST 08-MAR-2002
 DEFINITION BJ364861 Dictyostelium discoideum cDNA library, Cf Dictyostelium
 discoideum cDNA clone ddc33c11 5', mRNA sequence.
 ACCESSION BJ364861

VERSION BJ364861.1 GI:19274161
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum.
 Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 1 (bases 1 to 407)
 REFERENCES Urushihara, H., Tanaka, Y., Kohara, Y., and Shin-i, T.
 Full length cDNA of Dictyostelium discoideum at the culmination
 stage
 JOURNAL Unpublished (2002)
 CONTACT Tadasu Shin-i
 CENTER For Genetic Resource Information
 NATIONAL Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 TEL: 81-559-81-6856
 FAX: 81-559-81-6855
 EMAIL: tshin@genes.nig.ac.jp.

FEATURES
 source
 Location/Qualifiers
 1..407
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="ddc33c11"
 /clone_lib="Dictyostelium discoideum cDNA library, Cf"
 /sex="mat A"
 /dev_stage="Culmination stage"

BASE COUNT 203 a 64 c 29 g 108 t 3 others

ORIGIN

Query Match 8.6%; Score 41.2; DB 13; Length 407;
 Best Local Similarity 48.4%; Pred. No. 1.4; Mismatches 116; Indels 0; Gaps 0;
 Matches 109; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 75 AATPAGGATGTTGTTTAACCTTAAGTCAAAAAATPAGCAATTTCCACAGCAAAAGC 134
 Db 405 AATCATGTTGTTGTTAACTTACTAGATTAATTAATTAATTAATTAATTAATTAATTA 346
 QY 135 TAAATATCGCAAAACCTCGTAATAAATCTGCGAGATTAATGCTAAGTAACATTT 194
 Db 345 TCTTAATCTTCAAAATTAACAGTGAATTTGTTGATGTTATGATGATGATGATGATGA 286
 QY 195 ACCCCGTAAGTAAATGATCAATCAAGAGATGCGGCTGAATGATCGTCTATTG 254
 Db 285 ATTAAATGAAGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 226
 QY 255 AATTACAGCTGGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 299
 Db 225 TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 181

RESULT 6
 LOCUS AV867201 501 bp mRNA linear EST 08-NOV-2001
 DEFINITION AV867201 Nori Satoh unpublished cDNA library, egg Clona
 intestinalis cDNA clone c1eg50106 3', mRNA sequence.
 ACCESSION AV867201
 VERSION AV867201.1 GI:16654725
 KEYWORDS EST.
 SOURCE Clona intestinalis.
 ORGANISM Clona intestinalis.
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Clona.
 1 (bases 1 to 501)

REFERENCE
 AUTHORS Satoh, N., Satou, Y., Kohara, Y., and Shin-i, T.
 JOURNAL Unpublished genes in Clona intestinalis
 COMMENT Unpublished (2000)
 CONTACT Nori Satoh
 DEPARTMENT of Zoology
 KYOTO University
 Sakyo-ku, Kyoto 606-8502, Japan
 TEL: 81-75-753-4081
 FAX: 81-75-705-1113
 EMAIL: satoh@ascidian.zool.kyoto-u.ac.jp.

| | |
|---|--|
| FEATURES | location/Qualifiers |
| SOURCE | 1..501 /organism="Clona intestinalis" /db_xref="taxon:7719" /clone="cleg50106" /clone_lib="Nori Satoh unpublished cDNA library, egg" /tissue_type="whole animal" /dev_stage="egg" /note="Vector: pbluescript SK" |
| BASE COUNT | 177 a 65 c 88 g 163 t 8 others |
| ORIGIN | |
| Query Match | 8.4%; Score 40.2; DB 10; Length 501; |
| Best Local Similarity | 47.8%; Pred. No. 2.6; |
| Matches 117; Conservative 0; Mismatches 128; Indels 0; Gaps 0; | |
| OY | 9 ATTGAATGATGCTATTGGCTATTGTACACTTTCCTTTACCCAGCGTCAACAGTTTAATAACT 68 |
| DB | 150 ATAAAGTAAAGTAAAGTAAAGTAAATATACACGATGTTGCACATCTGGCAGATTAAAGTCT 209 |
| OY | 69 GCCAGCATTAAGGGATGTTGTTTACTTAAGTCAAAATAATGCCAATTTTCCACGAAAC 128 |
| DB | 210 GGCACTATACACGAATAATTTGCAATTTATTAACCAATATACCAATTTACAAAAATGAAA 269 |
| OY | 129 AAAAGCTAAATATATGCCAAAACCAGTAAATCTTGCTGGAGCTAATATTTGTAAGTA 188 |
| DB | 270 AAATCTTAAATATTTGTACAAATGTGTGCATTTATTCGATTTGAATGTTTCAGTTCAAAGAA 329 |
| OY | 169 ACATTTACCCCTGTAAGTTAATGATTCATATCAAGACAGATGGGCTGAATGAATCGTC 248 |
| DB | 330 AGTGAGAAGTCAGAAATATGTTCAATTAATTTAAAGAAAATATATATGTTGATGTAC 389 |
| OY | 249 TTATT 253 |
| DB | 390 GTCTT 394 |
| RESULT 7 | |
| BH146288 | |
| LOCUS | BH146288 940 bp DNA linear GSS 27-AUG-2001 |
| DEFINITION | ENTPND9TF Entamoeba histolytica sheared DNA Entamoeba histolytica genomic, DNA sequence. |
| ACCESSION | BH146288 |
| VERSION | BH146288.1 GI:15301168 |
| KEYWORDS | GSS. |
| SOURCE | Entamoeba histolytica. |
| ORGANISM | Entamoeba histolytica. Eukaryota; Eutromebidae; Entamoeba. |
| REFERENCE | 1 (bases 1 to 940) Loftus,B., Wang,Z., Van Aken,S. and Fraser,C. Determination of clone end sequences from Entamoeba histolytica HMI:IMSS sheared DNA library (2001) unpublished |
| AUTHORS | Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 |
| JOURNAL | Email: b.loftus@igr.org |
| COMMENT | Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library Seg primer: M13-Forward Class: shotgun High quality sequence start: 15 High quality sequence stop: 703. Location/Qualifiers 1..940 /organism="Entamoeba histolytica" /strain="HMI:IMSS" /db_xref="taxon:5759" /clone_lib="Entamoeba histolytica sheared DNA" /note="Vector: pHD51; Site_1: Bst I; Constructed at The |
| FEATURES | |
| SOURCE | |

| | | | | | |
|-----------------------|------------------------------------|------------------|---|--|-----|
| BASE COUNT | 345 a | 89 c | 114 g | 392 t | |
| ORIGIN | | | | | |
| Query Match | 8.1%; Score 39; DB 17; Length 940; | | | | |
| Best Local Similarity | 49.7%; Pred. No. 5.7; | | | | |
| Matches | 99; Conservative | 0; Mismatches | 100; Indels | 0; Gaps | |
| 0; | 0; | | | | |
| QY | 81 | GGAGTTGCTTAAC | TAGTC | CAAAAATAGCAATTTTCCACGACAAAGTAATA | 140 |
| DB | 595 | GTAATTTCTTTCAC | CAAAAAAAGAACCAATGATTTTTCATTAATATGGATAATTG | 654 | |
| QY | 141 | TCGC | AAAAACCTCAGTA | AAAAATCTTGCTGAGACTATTAATGCTAAGTAACAATTNACCCC | 200 |
| DB | 655 | TTTCATTTAAN | CAATATTTATTTATTTATTTGACATATTTATTCGCTTTTATTTATTTAAACG | 714 | |
| QY | 201 | TGAAGTAA | TGAGATCAATCAACAGAGAGATGGCGCTGTAATGCAATCGTCTTATTTGAATTAA | 260 | |
| DB | 715 | TTTCATTCATGATGAT | ATAAATAATATATGCTAATAATATATATACAAAAACAATATGGAATGGA | 774 | |
| QY | 261 | CAGGTGAGTGC | GTCTGT | 279 | |
| DB | 775 | TATTTTCTCTCTTTT | TGCT | 793 | |

| | |
|------------|--|
| RESULT | 8 |
| LOCUS | AZ347477 663 bp DNA linear GSS 29-SEP-2000 |
| DEFINITION | 1M00830005 Mouse 10kb plasmid UUCU1M library Mus musculus genomic clone UUCU1M0083005 R, DNA sequence. |
| ACCESSION | AZ347477 |
| VERSION | AZ347477.1 GI:10426714 |
| KEYWORDS | GSS. |
| SOURCE | house mouse. |
| ORGANISM | Mus musculus |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathii; Muridae; Murinae; Mus. 1 (bases 1 to 663) |
| AUTHORS | Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R. |
| TITLE | Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts |
| JOURNAL | Unpublished (2000) |
| COMMENT | Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 306, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunne@genetics.utah.edu Insert Length: 10000 Std Error: 0..00 Plate: 0083 row: O column: 05 Seq primer: CACACAGAACACGTATGACC Class: plasmid ends High quality sequence stop: 663. Location/Qualifiers 1..663 /organism="Mus musculus" /strain="C57BL/6J" |
| FEATURES | |
| Source | |

| | | | |
|--------------------------|---|---------------------|--------------------|
| FEATURES | Plate: ILCM266 | row: e | column: 04 |
| | High quality | sequence stop: 775. | |
| source | Location/Qualifiers | | |
| | 1. .891 | | |
| | /organism="Homo sapiens" | | |
| | /db_xref="taxon:9606" | | |
| | /clone="IMAGE:3610347" | | |
| | /clone_lib="NIH MGC 39" | | |
| | /tissue_type="adenocarcinoma" | | |
| | /lab_host="DH10B (phage-resistant)" | | |
| | /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." | | |
| BASE COUNT | 162 a | 277 c | 292 g 160 t |
| ORIGIN | | | |
| Query Match | 7.9%; | Score 38.2; | DB 10; Length 891; |
| Best Local Similarity | 54.7%; | Pred. No. 9.3; | |
| Matches 76; Conservative | 0; | Mismatches 63; | Indels 0; Gaps 0; |
| Y 245 | CGCTTATTGATTAACAGGTGATCGCTTCTTGCGTTTCAGTCATCTCTTGCGCGT | 304 | |

| | | | | | |
|----|---|--|-----|--|-----|
| D | b | | 746 | CCTTCATTGAGGGTCCCGAGAAGATTTGATCTCATCATTCAAGGCATCCACCTTGCCCTC | 687 |
| OY | | | 305 | GCGAGTCACAATTGAACAATAAACCACACTTAACAGAGTGTCTGGTAGAACACAACTAA | 364 |
| | | | | | |
| D | b | | 686 | CAGCTCCACCTTGGTCATATTAGCGAGCATCCACATCTCTTCAGCACACAAAATCAT | 627 |
| OY | | | 365 | GCTCTGCACCTTGTGGAGCG | 383 |
| | | | | | |
| D | b | | 626 | TCTCAGCAGCTGTGCGGCG | 608 |

RESULT 10

| | | | | |
|------------|--|--|--------|-----------------------------|
| Locus | H90453 | | 452 bp | mRNA linear EST 28-NOV-1995 |
| Definition | vz01c05.r1 Soares fetal liver spleen lINfLS Homo sapiens cDNA clone IMAGE:241448 5' , mRNA sequence. | | | |
| Accession | H90453 | | | |
| Version | H90453.1 | | | |
| Keywords | EST . GI:1080883 | | | |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.
 1 (bases 1 to 452)
 Haller, E., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
 M., Hulman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Riklin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston
 R., Williamson, A., Wohldmann, P. and Wilson, R.
 The WashU-Merck EST Project
 Unpublished (1995)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 2214
 High quality sequence stops: 210
 Source: IMAGE Consortium, LINL
 This clone is available royalty-free through LINL ; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 2214 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 210.
 Location/Qualifiers
 1. 452

```

/organism="Homo sapiens"
/db_xref="GDB:3790581"
/db_xref="taxon:9606"
/clone="IMAGE:241448"
/clone_id="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pTV73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACCTGACAGATTAATTAAAGACTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pTV73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bernaldo."

```

[illegible]

| | |
|------------|--|
| RESULT | 11 |
| BJ370199 | |
| LOCUS | BJ370199 633 bp mRNA linear EST_08-MAR-2002 |
| DEFINITION | BJ370199 Dictyostelium discoideum cDNA library; CF Dictyostelium discoideum cDNA clone ddc53c12 5' , mRNA sequence. |
| ACCESSION | BJ370199 |
| VERSION | BJ370199.1 GI:19279582 |
| KEYWORDS | EST. |
| SOURCE | Dictyostelium discoideum . Dictyostelium discoideum Eukaryota; Mycetozoa; Dictyosellidae; Dictyostelium. 1 (bases 1 to 633) Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T. Full length cdna of Dictyostelium discoideum at the culmination stage |
| ORGANISM | |
| REFERENCE | |
| AUTHORS | |
| TITLE | |

| | |
|-----------------|--|
| JOURNAL COMMENT | Unpublished (2002) Contact: Tadasu Shin-1 Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshin@genes.nig.ac.jp. |
| FEATURES | Location/Qualifiers |
| source | 1. 633 /organism="Dictyostelium discoideum" /strain="AX4" /db_xref="taxon:44689" /clone="ddc53c12" /clone_lib="Dictyostelium discoideum cDNA library, CF" /sex="mat A" /dev_stage="Culmination stage" |
| BASE COUNT | 226 a 73 c 116 g 217 t 1 others |
| ORIGIN | |
| Query Match | 7.9%: Score 38; DB 13; Length 633; |

| | Best Local Similarity | 47.8% | Pred. No. 10; | Matches 110; | Conservative 0; | Mismatches 120; | Indels 0; | Gaps 0; |
|----|-----------------------|--|---------------|--------------|-----------------|-----------------|-----------|---------|
| QY | 73 | GCATTAAGGAGTGTGTTAACTTAATGCAAAAAAATATGCGAATTTTCCACGACAA | 132 | | | | | |
| Db | 106 | GCACCAAGAGAAAGATGTTGATGTATTAATATGTTATGCGAATTTGATTAATATACCA | 165 | | | | | |
| QY | 133 | GCTAATATATGCGAAAAACCTCAGTAAAAATCTTGCTGGAGCTATTAATTGCTAATGA | 192 | | | | | |
| Db | 166 | GGTGTAAATTCACATAATTAATAGTAATTAATGAGCAAGAGATGATGATCATCT | 225 | | | | | |
| QY | 193 | TTACCCCTCGAAGTTAATGATCAATCAAGAGATGTGGCGCTGAATGAATCGTCTTAT | 252 | | | | | |
| Db | 226 | TATACAACAAAATGTTGATGATGTATGTTAATTTAAATTTTTCACAAATATGTAAT | 285 | | | | | |
| QY | 253 | TGAATTAACAGTGGATCGTCTTGTTGGTTTCACTCAATTCCTCTTGGGG | 302 | | | | | |
| Db | 286 | TTAATAAAATTAATTTGGCAGAGATCAAGTGACACATATTTTATAGTGGTTGGGG | 335 | | | | | |

| | RESULT | 12 |
|------------|--|--|
| LOCUS | BH161244 | |
| DEFINITION | BH161244 | 864 bp DNA linear GSS 24-SEP-2001 |
| ACCESSION | ENTRX81ITF Entamoeba histolytica Sheared DNA | Entamoeba histolytica genomic, DNA sequence. |
| VERSION | BH161244 | |
| KEYWORDS | BH161244.1 GI:15734682 | GSS. |
| SOURCE | Entamoeba histolytica. | |
| ORGANISM | Entamoeba histolytica. | |
| REFERENCE | Eukaryota; Entamoebidae; Entamoeba. | |
| AUTHORS | 1 (bases 1 to 864) | |
| TITLE | Lofthus,B., Wang,Z., Van Aken,S. and Fraser,C. | |
| JOURNAL | Determination of clone end sequences from Entamoeba histolytica | |
| COMMENT | HMI:IMSS sheared DNA library (2001) | |
| | Unpublished (2001) | |
| | Contact: Brendan J Lofthus | |
| | Department of Eukaryotic Genomics | |
| | The Institute for Genomic Research | |
| | 9712 Medical Center Dr., Rockville, MD 20850, USA | |
| | Tel: 301 838 0208 | |
| | Fax: 301 838 3543 | |
| | Email: b.j.lofthus@tigr.org | |
| | Clones are derived from the Entamoeba histolytica HMI:IMSS sheared | |
| | DNA library | |
| | Seq primer: M13-Forward | |
| | Class: shotgun | |
| | High quality sequence start: 5 | |
| | High quality sequence stop:728. | |

| FEATURES | SOURCE | Location/Qualifiers |
|-----------------------|--------|--|
| | | 1. 864 |
| | | /organism="Entamoeba histolytica" |
| | | /strain="HML:IMSS" |
| | | /db_xref="taxon:5759" |
| | | /clone_lib="Entamoeba histolytica Sheared DNA" |
| | | /note="Vector: pHOS1; Site: 1; Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barrell, Oxford University Press, 1999)." |
| BASE COUNT | 341 a | 94 c 99 g 330 t |
| ORIGIN | | |
| Query Match | 7.9% | Score 38; DB 17; Length 864; |
| Best Local Similarity | 48.2% | Pred No. 11; |

